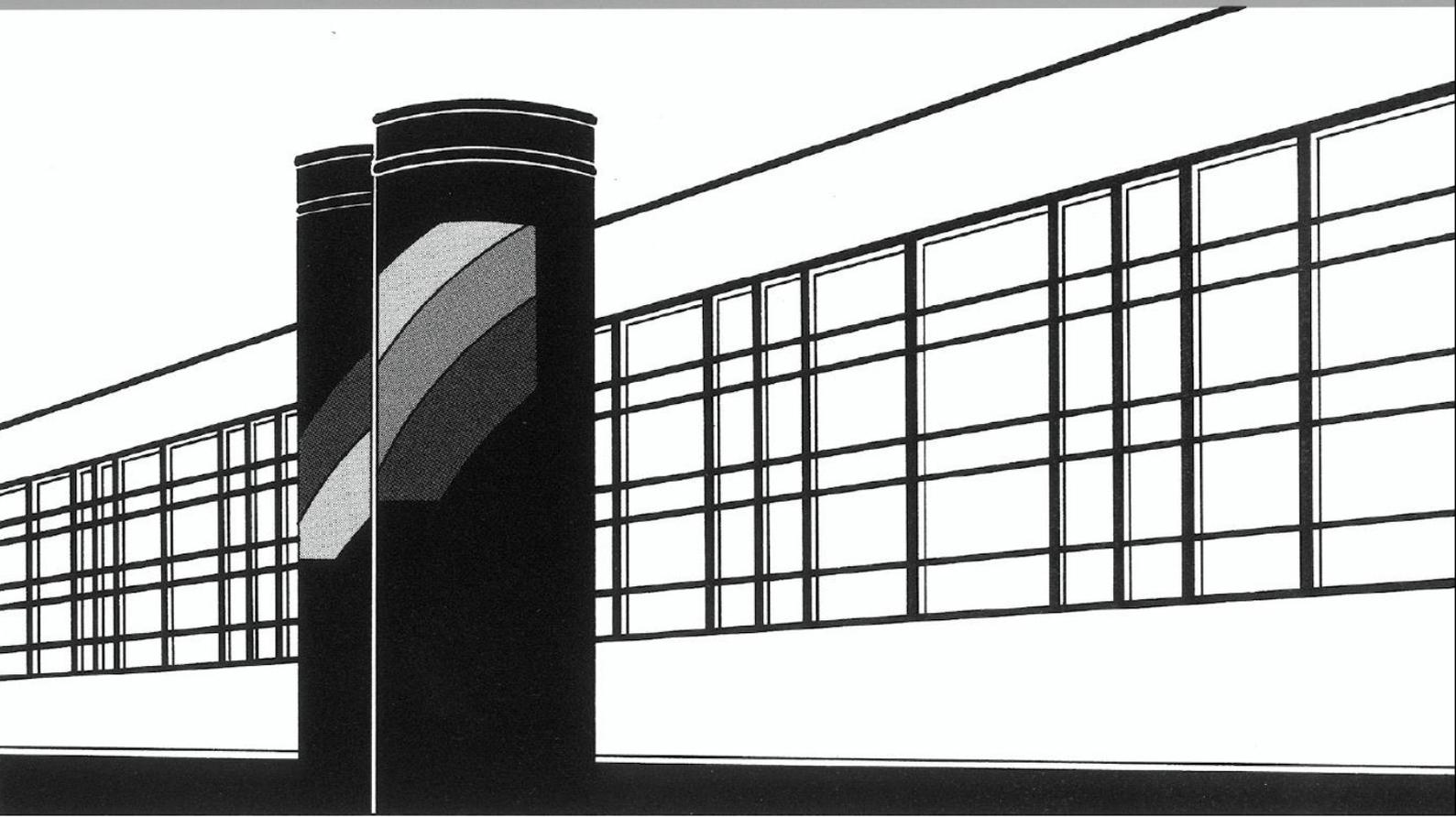


Universität Stuttgart



Institut für Wasser- und Umweltsystemmodellierung

Mitteilungen



Heft 252 Thomas Müller

Generation of a Realistic Temporal Structure
of Synthetic Precipitation Time Series for
Sewer Applications

Generation of a Realistic Temporal Structure of Synthetic Precipitation Time Series for Sewer Applications

von der Fakultät Bau- und Umweltingenieurwissenschaften der
Universität Stuttgart zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.) genehmigte Abhandlung

vorgelegt von
Thomas Müller
aus Karlsruhe, Deutschland

Hauptberichter: Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy

Mitberichter: Prof. Dr.-Ing. Uwe Haberlandt

Prof. Dr.-Ing. Heidrun Steinmetz

Tag der mündlichen Prüfung: 06. November 2017

Institut für Wasser- und Umweltsystemmodellierung
der Universität Stuttgart
2017

Heft 252 **Generation of a Realistic
Temporal Structure of
Synthetic Precipitation Time
Series for Sewer Applications**

von
Dr. -Ing.
Thomas Müller

D93 Generation of a Realistic Temporal Structure of Synthetic Precipitation Time Series for Sewer Applications

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://www.d-nb.de> abrufbar

Müller, Thomas:

Generation of a Realistic Temporal Structure of Synthetic Precipitation Time Series for Sewer Applications, Universität Stuttgart. - Stuttgart: Institut für Wasser- und Umweltsystemmodellierung, 2017

(Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart: H. 252)

Zugl.: Stuttgart, Univ., Diss., 2017

ISBN 978-3-942036-56-6

NE: Institut für Wasser- und Umweltsystemmodellierung <Stuttgart>: Mitteilungen

Gegen Vervielfältigung und Übersetzung bestehen keine Einwände, es wird lediglich um Quellenangabe gebeten.

Herausgegeben 2017 vom Eigenverlag des Instituts für Wasser- und Umweltsystemmodellierung

Druck: Document Center S. Kästl, Ostfildern

Danksagung

... and finally it's done!

Allerdings hätte ich die vorliegende Arbeit ohne die Unterstützung, die Ratschläge und Ermutigung vieler Personen aus meinem Umfeld nicht beendet, ja, wahrscheinlich noch nicht einmal begonnen. Daher möchte ich mich bei allen Wegbegleitern bedanken, die mir in den letzten Jahren mit unterschiedlichster Hilfe innerhalb der Projekte, während der Vorlesungen und Übungen, im Arbeitsalltag oder privat zur Seite standen.

Einige Personen möchte ich speziell hervorheben. Allen voran, Herrn Prof. Bárdossy, dass Sie mich nicht nur als Doktorand annahmen, sondern mich stets sowohl mit kreativen, neuen Ideen als auch mit Lösungsvorschlägen zu auftretenden Problemen unterstützten. Des Weiteren möchte ich Frau Prof. Steinmetz und Herrn Prof. Haberlandt danken, dass sie sich als Mitberichter zur Verfügung stellten, sowie Herrn Prof. Kranert für die Übernahme des Prüfungsvorsitzes und Gabriele Hartmann für die Organisation des Doktorandenprogramms ENWAT.

Besonders möchte ich dir, Jochen, danken, dass du es ermöglicht hast, in einem lockeren Arbeitsumfeld zu arbeiten, das wenigen bürokratischen Zwängen unterlag; ebenso dir, Astrid, dass du nicht nur dabei halfst, das darüber hinaus aufgetretene bürokratische Dickicht zu lichten, sondern mir auch bei sprachlichen Unklarheiten motiviert unter die Arme gegriffen hast.

Meinen Dank möchte ich ebenso allen weiteren Kollegen des IWS aussprechen, insbesondere Dirk für deine Hilfen beim Programmieren und Sinnieren über Hypothesentests; Felix für deine humorigen Kommentare und ernsthaften Gespräche in den Pausen; an Basti für die gemeinsame Organisation und Vorbereitung vieler Statistik-Übungen; an Micha für deine Zusammenarbeit beim Programmieren der Webseite der Wetterstation; an Faizan für deine augenöffnende Sichtweise auf Deutschland und die Welt; an Julia und Michael für eure schnelle und unkomplizierte Hilfe bei Computerproblemen und vor allem an dich, „Twin“ Tobi, für deine aufbauenden Worte bei Rückschlägen, deine Verlässlichkeit und schlicht und einfach die sehr gute kollegiale Zusammenarbeit. Hervorheben möchte ich außerdem Ferdi. Durch deine Übungen hast du mich zunächst für das Themengebiet Statistik interessiert, hast mich während der Diplomarbeit unterstützt und mir letztendlich mit dem erfolgreichen Antrag der Projekte SYNOPSE und SAMUWA den Weg zu dieser Promotion geebnet. An dieser Stelle möchte ich mich zudem beim BMBF für die Finanzierung dieser beiden Projekte bedanken.

Weiterhin möchte ich allen Projektteilnehmern für die gute Zusammenarbeit danken. Hervorheben möchte ich Kai Schröder von Pecher und Partner und Stefanie Maßmann vom

itwh für die zusätzlichen Schmutzfrachtsimulationen mit dem Kanalnetz Braunschweig und Freiburg, sowie Manfred Schütze vom ifak Magdeburg für die Simulationen mit Kanalnetzsteuerung. Ebenso möchte ich mich bei Uli Dittmer vom ISWA der Universität Stuttgart für die fachliche Unterstützung im Bereich der Siedlungsentwässerung bedanken.

Meinen Dank möchte ich außerdem meinen Bachelor-, Masterstudenten und Hiwis aussprechen (Ann-Kathrin, Antonio, David, Janina, Kanwal, Sarah und Tobi). Ihr habt mir einerseits viel Arbeit abgenommen und mich andererseits durch eure Ergebnisse zum weiteren Nachdenken und zum Untersuchen von diversen neuen Aspekten angeregt.

Zu guter Letzt möchte ich Lena, Theo, meinen Eltern, meiner Schwester und meinen Freunden bedanken, dass ihr mir stets vor Augen führt, dass es noch wichtige und wichtigere Dinge im Leben außerhalb der Uni gibt.

Stuttgart, November 2017

Thomas Müller

Contents

List of Figures	IV
List of Tables	VI
List of Abbreviations	VII
List of Notations	VIII
Summary	XI
Zusammenfassung	XVII
1. Introduction	1
2. Observed precipitation	3
2.1. Overview of measurement devices	3
2.2. Statistical characterization of precipitation	4
2.2.1. Global statistics	5
2.2.2. Event-based statistics	8
2.2.3. Extreme value statistics	8
2.2.4. Other approaches	10
3. Analysis of observed precipitation	11
3.1. Measurement errors and inhomogeneity	11
3.2. Overview of applied measured data	12
3.3. Influence of different measurement devices on precipitation statistics	15
3.4. Issues using scaling parameters	18
3.4.1. Breaks in scaling property	19
3.4.2. Sensitivity to measurement devices	19
3.5. Issues with extreme values	21
4. Synthetic precipitation	23
4.1. Motivation	23
4.2. Alternatives to measured time series	25
4.3. Overview of different precipitation models	26
4.4. The NiedSim approach	28
4.4.1. Database	29
4.4.2. 1 h time step generation	31
4.4.3. 5 min disaggregation	32

4.5. Validation of synthetic precipitation	32
5. Sewer systems	35
5.1. Overview of sewer systems	36
5.1.1. Hydraulic design of sewer networks	37
5.1.2. Design and simulation of combined sewer systems	38
5.1.3. Sewer system management	39
5.2. Sewer model representation	40
5.2.1. Runoff generation	42
5.2.2. Overland flow	43
5.2.3. Sewer flow	43
5.2.4. Surface flooding	45
6. Setup of an artificial sewer model for an indirect validation	46
6.1. Configuration and interpretation of the artificial sewer	47
6.2. Modifications of the original artificial sewer	49
6.3. Implementation of SWMM	50
6.3.1. Runoff generation	50
6.3.2. Overland flow	50
6.3.3. Sewer flow routing	51
6.3.4. Post processing	52
6.4. Sensitivity analysis of the artificial sewer	53
6.5. Validation of the artificial sewer	55
7. Preliminary investigations of the indirect validation using the artificial sewer	60
7.1. Identification of sampling uncertainty	61
7.1.1. Assessment of temporal uncertainty	61
7.1.2. Origins of temporal uncertainty	63
7.1.3. Assessment of spatial uncertainty	67
7.2. Important characteristics of precipitation with respect to combined sewers . .	68
7.2.1. Influence of the intra-hourly intensity structure	68
7.2.2. Influence of extreme values	70
7.2.3. Influence of small values	72
8. Analysis of the hourly optimization of NiedSim	75
8.1. Motivation	76
8.2. Modifications of the hourly optimization	77
8.2.1. Autocorrelation and scaling parameters	79
8.2.2. Positioning of large values in a Poisson process	82
8.2.3. Including wet spell durations	91
8.3. Final discussion on the hourly optimization	96
9. Analysis of the disaggregation of NiedSim	102
9.1. Motivation	102
9.2. Modifications of the disaggregation algorithm	103
9.2.1. Autocorrelation and scaling parameters	104

9.2.2. Entropy	112
9.3. Final discussion on the disaggregation	116
10. Additional properties of rainfall relevant for sewer systems	118
10.1. Theory of asymmetry of dependence	118
10.2. Data analysis using asymmetry of dependence	121
10.3. Influence of asymmetry of dependence on combined sewers	122
11. Conclusions and Outlook	124
Bibliography	i
A. Appendix	xi

List of Figures

3.1. Time series of different periods	13
3.2. EDFs of measured data	16
3.3. ACF of measured data	17
3.4. Influence of different aggregation levels on the scaling parameters	20
3.5. Partial series of measured data	22
6.1. CSO statistic of original artificial sewer of <i>Drechsel</i> (1991)	49
6.2. Sensitivity analysis of artificial sewer	54
6.3. NS vs. fraction of total CSO	57
7.1. Bootstrap bandwidth	62
7.2. Q-q-transformed precipitation	64
7.3. Q-q-transformed overflow characteristics	65
7.4. Uncertainty of marginals vs. structure	66
7.5. Marginals and structure mixed from different stations	67
7.6. Influence of manipulated measured time series	69
7.7. Influence of small hourly values	72
7.8. Influence of small 5 min values	74
8.1. CSO statistics of current state of hourly optimization	77
8.2. CSO statistics of different scaling versions of hourly optimization	80
8.3. Partial series of scaling and autocorrelation versions (hourly)	81
8.4. Fictive spell to show the characteristics of the autocorrelation	82
8.5. Frequencies of large values vs. precipitation sums	85
8.6. Fitted sine function to slope parameter	87
8.7. Partial series of versions with Poisson process	90
8.8. CSO statistics of versions with Poisson process	91
8.9. Histogram of wet spell durations	92
8.10. CSO statistics of hourly optimization including wet spells (Freiburg)	93
8.11. Histogram of wet spell durations	94
8.12. CSO statistics of hourly optimization including wet spells (Freiburg)	95
8.13. Precipitation spell statistics of different hourly optimizations	97
8.14. Multivariate distribution of precipitation spells I (Kempten)	99
8.15. Multivariate distribution of precipitation spells II (Kempten)	100
9.1. CSO statistics of disaggregation with parameters of different periods	103
9.2. CSO statistics of scaling and autocorrelation versions	107
9.3. Partial series of scaling and autocorrelation versions (sub-hourly)	109

9.4. Precipitation event statistics of scaling and autocorrelation versions I	110
9.5. Precipitation event statistics of scaling and autocorrelation versions II	111
9.6. Scatter plot of entropy vs. hourly precipitation depth	113
9.7. CSO statistics of entropy versions	114
9.8. Scatter plot hourly entropy vs. precipitation sum (5 min)	115
10.1. Original time series and reversed time series	119
10.2. Empirical copula density and visualized asymmetry	120
10.3. Asymmetry of Augsburg and Kempten	122
10.4. Deviations in CSO statistics caused by reversed time series	123
A.1. 3d-surface fitting of Poisson process	xi
A.2. Scatterplot elevation vs. monthly frequency of Poisson values	xii
A.3. Scatterplot elevation vs. normalized monthly frequency of Poisson values . .	xiii
A.4. Semivariogram of the Poisson values	xiv
A.5. Partial series of versions with Poisson process	xv
A.6. Multivariate distribution of precipitation spells (Freiburg)	xvi
A.7. Scatter plot hourly entropy vs. precipitation sum (10 and 15 min)	xvi

List of Tables

2.1. Precipitation statistics	5
3.1. Overview of the time periods used for different locations.	13
3.2. Overview of some basic global statistics	14
3.3. Overview of the characteristics of different periods	16
4.1. Parameters used in NiedSim2.x	30
6.1. Overview of artificial sewer system	48
6.2. Classification of combined sewer overflow (CSO) structures of Braunschweig with artificial sewer ($l_o = 33$ m)	58
6.3. Classification of CSO structures of Braunschweig with artificial sewer ($l_o = 33$ and 5000 m)	59
7.1. Acceptable bandwidth of deviations	63
8.1. Overview of different hourly generation configurations	78
8.2. Validation of scaling parameter and autocorrelation of different hourly generations	79
9.1. Validation of scaling parameter and autocorrelation of different disaggregations	106
9.2. Probability of zero precipitation of different disaggregations	109
9.3. Parameters used in NiedSim3	117

List of Abbreviations

AMDA	<i>Automatische Messdatenerfassungsanlage</i> [automated gauging station]
ACF	autocorrelation function
AR	autoregressive process
CDF	cumulative distribution function
COD	chemical oxygen demand
CP	circulation pattern
CSO	combined sewer overflow
DDF	depth-duration-frequency
DWD	<i>Deutscher Wetterdienst</i> [German Meteorological Service]
EDF	empirical distribution function
EPA	US Environmental Protection Agency
GEV	generalized extreme value
KOSTRA	<i>Koordinierte Starkniederschlagsregionalisierung – Auswertung</i> [analysis of the coordinated regionalization of intense precipitation]
IDF	intensity-duration-frequency
IID	independent and identically distributed
LID	low impact development
MC	Markov Chain
MCMC	Markov Chain Monte Carlo
NiedSim	<i>NiederschlagsSimulator</i>
NS	Nash-Sutcliffe efficiency coefficient
PDF	probability density function
POT	peak over threshold
RCT	real-time control
SCS	Soil Conservation Service
SUDS	sustainable urban drainage systems
SWMM	Storm Water Management Model
USDA	United States Department of Agriculture
WSUD	water sensitive urban design
WWTP	wastewater treatment plant
spell	Consecutive precipitation values simply separated by at least one dry hour
event	Consecutive precipitation values separated by a sophisticated approach

List of Notations

The most important notations are listed below; less relevant symbols are explained in the respective section where they occur.

Latin letters

a_m	interception of linear regression between $n_{poiss,m}$ and H_m of month m
b_k	scaling parameter of moment k
b_m	slope of linear regression between $n_{poiss,m}$ and H_m
d_e	minimum event depth applied for precipitation event definition
$H^{(j)}$	precipitation time series at aggregation j
$H_{wet}^{(j)}$	precipitation time series $H^{(j)} > 0$ at aggregation j
H_S	Shanon entropy
i_e	minimum 5 min intensity applied for precipitation event definition
l_o	length of overland flow path
m_k^j	moment k at aggregation j
n_{nan}	number of missing data in a measured time series
n_{poiss}	number of values predefined in a Poisson process
\bar{n}_{poiss}	mean number of values of the Poisson process defined by x_{poiss}
n_{poiss}^+	number of values of the extended set of the Poisson process
$n_{wet}^{(j)}$	number of precipitation values $H^{(j)} > 0$ at aggregation j
\mathcal{N}	normal distribution
$p_e^{(j)}$	probability of exceeding a given threshold at aggregation j : $H^{(j)} > H_{thresh}$
p_I	probability that the values of the Poisson process are independent
Q_d	throttle discharge
Q_m	combined sewage discharge to the wastewater treatment plant (WWTP)
q_r	runoff discharge rate
$r_l^{(j)}$	autocorrelation of lag l at aggregation j
s_x	standard deviation
S_j	aggregations used for the calculation of the scaling parameters
S_{thresh}	threshold aggregation used to define a subset S_j
t_e	calculated emptying time
ts_e	separation time applied for precipitation event definition
V_s	specific storage volume
\bar{x}	mean
$ws_n^{(h)}$	hourly wet spell frequency of length n

$ws_{\Sigma}^{(h)}$	sum of all hourly wet spell frequencies
x_{poiss}	precipitation threshold which defines \bar{n}_{poiss} of the Poisson process

Greek letters

$\gamma(h)$	variogram
λ	parameter of the Poisson process
ϕ	statistics describing the marginals of precipitation values
ψ	statistics describing the temporal structure of precipitation values
ω	weighting factor of the parameters in the objective function

Summary

The collection and disposal of wastewater and stormwater with sewers are one of the most important parts of urban well-being. As the construction and maintenance of a sewer is very expensive, the simulation of its performance is a major part of the planning to assure the operational capability for the coming decades. An important input parameter of sewer simulations is rainfall. Because future precipitation is obviously not available, one assumes that measured rainfall time series from the past are representative of the future. Unfortunately, measured long-term series in high resolution are only available for a few locations. The generation of synthetic rainfall series can be a solution to substitute missing observations or to simulate scenarios for a future under changed conditions.

The stochastic precipitation generator developed at the University of Stuttgart is called *NiederschlagsSimulator* (NiedSim). The synthetic rainfall of the operational version of NiedSim is able to reflect the extremes and other statistics quite well, but shows deficiencies when the data is applied to combined sewer models. In this study, the algorithm of NiedSim is developed further with focus on a realistic temporal structure of the time series. A deep understanding of the characteristics of observed rainfall is fundamental for the enhancement of NiedSim as well as their influence on sewer simulations. Therefore, the analysis of measured data is an integral part of this study in addition to the actual development of NiedSim.

In the study, the measured data of four different locations are analyzed (Freiburg, Kempten, Weiden and Augsburg) reflecting different climates in Southern Germany. A comparison of statistics of time periods with varying measurement devices exhibit systematic deviations for high resolution data (< 1 h). Three time periods can be defined with homogeneous characteristics: (I) before 1990s, digitalized paper records with a smoothed 5 min aggregation (resolution 0.01 mm); (II) 1990s to early 2000s, discrete pulses of tipping buckets with a rather coarse resolution (0.1 mm); (III) since late 2000s, digital devices with the most realistic structure of a high resolution (< 0.1 mm). The data of (I) shows reduced p_0 , increased scaling parameters and autocorrelations, large deviations in the empirical distribution function (EDF) of small intensities and changed event based statistics. There is also some evidence that even extreme values at a high resolution (5 min) could be influenced.

The objective of NiedSim is to generate realistic and robust synthetic time series for sewer applications. Depending on the application, different characteristics of the rainfall input are relevant. For the hydraulic design of sewer networks, large intensities are crucial that can be reflected well by partial series, whereas for the planning of combined sewer systems, long-term mean precipitation characteristics are important.

A validation method for synthetic time series is an indirect validation using sewer simulations. The artificial sewer of this study is based on a modified and extended approach of

the model developed by *Drechsel* (1991) and represents realistic sewers in practice according to *ATV-A128* (1992). Not only the specific volume and specific discharge of overflow structures (with and without volume) are varied, but also the length of the overland flow path. The artificial sewer is implemented in the model *Storm Water Management Model (SWMM)* with an overland flow calculated with the continuity and Manning equations and a sewer flow computed with hydrological routing. Two major setups are developed with different lengths of overland flow paths representing very fast responding sewers as well as systems with a highly attenuated discharge.

A comparison of the mean yearly CSO statistics of the sewer in Braunschweig with the artificial sewer indicates that both setups are necessary to reflect different, real case systems. The absolute CSO output using both setups can represent well the various sewers in Baden-Württemberg specified by a statistical survey of *Brunner* (1997).

General conclusions between rainfall statistics and combined sewer overflows

The artificial sewer is extensively studied using measured precipitation data to assess the temporal and spatial uncertainty of rainfall as well as to define confidence intervals of CSO statistics serving as benchmarks for synthetic time series. The temporal sampling uncertainty is investigated by bootstrap methods. The results confirm the issues of short time periods (< 10 years) reported in literature. However, even for much longer periods (37 years) uncertainties in CSO statistics of up to 10 % can still occur.

Furthermore, the study shows with a q-q-transformation that the high yearly temporal sampling uncertainty can rather be referred to large yearly variations in the hourly marginal distributions and less to a varying temporal structure. The influence of the hourly marginals dominates the temporal order of the values with respect to CSOs, but only if the structure is realistic. Otherwise, large deviations can occur even for identical marginals.

If the marginals of one location are q-q-transformed on the structure of another location, large systematic CSO statistics can be observed. The deviations of the hourly temporal structure of different locations is larger compared to their individual variations among the years. However, the spatial variability of the structure is smaller compared to the temporal sampling uncertainty at one location if the marginals are included and short time periods are used.

The intra-hourly intensity structure is mainly important for CSO structures without volume of fast responding sewers. Otherwise, the attenuation and retention processes smooth the discharge series and equalize the intra-hourly intensity structure of the input data.

In many combined sewers the partial series, that is, the extremes are less important compared to other events not reflected by intensity-duration-frequency (IDF)-curves. More precisely, partial series have only a substantial influence if long durations (> 1 h) are considered, whereas for large intra-hourly intensities, their respective cumulative precipitation sum is too small for a significant contribution to CSOs.

Precipitation data is highly skewed and only around 20 % of all hourly values are larger than 1.0 mm, but provide approximately 70 % to the total precipitation sum. Nevertheless,

neglecting hourly values (< 1.0 mm) or small 5 min values (< 0.2 mm, depending on the measurement device) can have large impacts on CSO statistics. The reason is that the cumulative sum of several small values can contribute to the discharge if they are part of a larger precipitation event. Therefore, higher thresholds can only be applied if event based statistics are used, or if large aggregations are considered (e.g. daily resolution). For global statistics at high resolutions as well as for NiedSim, small values are important and, consequently, a threshold must be selected carefully.

The CSO volume is the most important parameter in combined sewer simulations in practice. As shown in the study, it is more robust to errors in the rainfall time series as compared to the CSO duration. A validation of the CSO duration can highlight additional deficiencies in the temporal structure because the duration can also be influenced by small rainfall values. Such issues are not only important if the CSO duration is of interest, but might also be relevant for other simulations not explicitly validated in the study.

Finally, a new measure, the *asymmetry of dependence*, is developed that demonstrates a temporal irreversibility of precipitation time series. The irreversibility affects the flow characteristics and, thus, CSO statistics of combined sewer simulations. The influence is in the same order as compared to the deviations of synthetic NiedSim series from measured series.

Performance of hourly optimization of NiedSim

The performance of hourly optimization of NiedSim is evaluated directly by a comparison of the statistics of the simulated time series with the target input statistics. Additionally, the data is validated indirectly using partial series, the artificial sewer and event based statistics. In order to avoid any influence of the hourly marginal distribution, measured time series with a randomized hourly structure are reconstructed with NiedSim without explicitly simulating the marginals. Large systematic deviations can be observed in the indirect validations using the operational version of NiedSim although the target statistics can be optimized in most cases accurately. The results emphasize the high complexity of rainfall-runoff-overflow processes that can neither be described by a few rainfall statistics, nor by considering only partial series, nor by simulating just a few overflow structures.

The autocorrelation and weighted autocorrelation in NiedSim lead to an indirect optimization of the scaling characteristics. Some deviations with respect to the scaling parameters are observable, but they cannot be reduced further if the scaling parameters are explicitly optimized. Issues with scaling parameters are shown (breaks in the scaling behavior and a dependence on the measurement device), and hence, they are not used anymore for the hourly optimization.

The optimization of the autocorrelation causes a strong clustering of large precipitation values. Therefore, a Poisson process is introduced that positions large hourly values fixed in the time series before the optimization. The number of fixed values per month depends on the season and on the monthly precipitation sum. In order to account for several large values occurring in one spell, not all largest values are prepositioned. The parameters of

the Poisson process are fairly robust with respect to the CSO statistics. As they are additionally independent of location and elevation, one set of Poisson parameters can be fitted for Southern Germany. The introduction of the Poisson process reduces the deviations of the time series of large values contributing to the partial series, and improves the entire temporal structure indicated by reduced errors in the CSO statistics.

Further issues of the synthetic series can be related to an erroneous intermittency. Too many very short precipitation spells with large volumes are generated, whereas the depth of longer spells is highly underestimated. Consequently, the wet spell frequency distribution is added as an additional target in the optimization. The validation of event statistics shows an improvement of the temporal structure, although a positive influence on the CSO statistics is not distinct for all locations.

The introduction of the Poisson process as well as the wet spell frequency distribution improves the temporal structure of the time series compared to the previous version of NiedSim. The remaining deviations from the measured time series in the partial series and in the CSO volume are small compared to the sampling uncertainty. However, a systematic underestimation of the CSO duration can still be observed depending on a particular sewer setup. Unfortunately, it is not possible to relate the remaining issues to further rainfall statistics like distribution functions or bivariate event-based statistics.

Performance of 5 min disaggregation of NiedSim

The study shows several challenges with respect to the disaggregation of hourly data. First of all, the influence of different measurement devices on statistics of high resolution data cannot be neglected. Consequently, only data of the most recent pluviometers (e.g. weighting devices) should be used to obtain homogeneous input statistics. Such time series show an increased sampling uncertainty due to a reduced length of available data for many locations. As the mean yearly input statistics of the NiedSim disaggregation are relatively robust, the effect of the increased sampling error could be tolerable. However, the larger uncertainty highly affects the long-term validation using sewer models and partial series and makes the evaluation difficult.

Other issues arise with the actual reconstruction of high resolution temporal structures. Already the performance of the individual 5 min optimization of autocorrelation and scaling parameters is worse compared to the hourly optimization. If both are optimized, the performance is further reduced and depends on the chosen weighting factor of the objective function. The best results are obtained using scaling parameters as well as weighted autocorrelation with a focus on the latter one in the optimization.

However, systematic issues in the temporal structure are observed. Too few zero values at high resolution time scales indicate that the hourly rainfall depth is spread across all 5 min values instead of forming spells with well defined starts and endings. An additional optimization on the distribution of the hourly entropy cannot improve the 5 min structure, but illustrates general difficulties in the 5 min optimization using the NiedSim approach.

Nevertheless, the NiedSim disaggregation improves the time series compared to a time series with randomly changed 5 min values. Furthermore, it is demonstrated that the partial series could be adjusted in a post-processing step without large impacts on CSO statistics.

In conclusion, as the issues in the disaggregation are less relevant for many combined sewer systems as compared to the rather well-performing hourly optimization, and as the large intensities of the partial series could be adjusted in a post-processing step, the optimization of the temporal structure of NiedSim is supposed to be applicable for most sewer simulations.

Zusammenfassung

Die Kanalnetze zum Sammeln und zum Transport von Abwasser und Regenwasser tragen wesentlich zum städtischen Wohlbefinden bei. Da die Bau- und Unterhaltskosten der Kanalnetze sehr hoch sind, werden in der Planungsphase Langzeitsimulationen durchgeführt, um das Leistungsvermögen der Netze über Jahrzehnte sicherzustellen. Ein entscheidender Eingangsparameter dieser Modelle ist Niederschlag. Da zukünftige Niederschlagszeitreihen offensichtlich nicht verfügbar sind, nimmt man an, dass der gemessene Niederschlag der Vergangenheit auch für die Zukunft repräsentativ ist. Bedauerlicherweise stehen gemessene, langjährige Niederschlagszeitreihen mit einer hohen zeitlichen Auflösung nur an wenigen Orten zur Verfügung. Eine Möglichkeit, um fehlende Messdaten zu ersetzen oder um Zukunftsszenarien mit unterschiedlichen Bedingungen zu simulieren, ist die Generierung von synthetischen Niederschlagszeitreihen.

Ein stochastischer Niederschlagsgenerator wurde an der Universität Stuttgart unter dem Namen *NiedSim* (NiederschlagsSimulator) entwickelt. Die synthetischen Niederschlagszeitreihen der operationellen Version *NiedSim* können Extremwerte sowie andere Statistiken verhältnismäßig gut wiedergeben. Es zeigen sich jedoch Mängel bei der Anwendung der Reihen in Schmutzfrachtmodellen. In dieser Studie wird *NiedSim* mit dem Ziel weiterentwickelt, die zeitliche Struktur der Zeitreihen möglichst realistisch abzubilden. Für die Entwicklung eines Niederschlagsmodells ist ein tiefes Verständnis der gemessenen Niederschlagseigenschaften in Bezug auf das Abflussverhalten im Kanalnetz wesentlich, weswegen die Untersuchung von beobachteten Zeitreihen ein großer Teil der Studie umfasst.

In der Studie werden Messdaten von vier verschiedenen Orten untersucht (Freiburg, Kempten, Weiden und Augsburg), welche Standorte in Süddeutschland mit unterschiedlichen Klimata abdecken. Ein Vergleich von Statistiken verschiedener Zeitperioden mit diversen Niederschlagsmessgeräten zeigt systematische Unterschiede bei hochaufgelösten Datensätzen (< 1 h). Drei Zeiträume mit homogenen Eigenschaften können bestimmt werden: (I) vor den 1990er Jahre, digitalisierte Niederschlagsschreiber mit einer gleichmäßigen fünfminütigen Intensität (Auflösung 0,01 mm); (II) 1990er bis frühe 2000er, Kippwaagen mit diskreten Impulsen in relativ grober Auflösung (0,1 mm); (III) seit den späten 2000er, digitale Messgeräte mit der realistischsten Struktur und einer hohen Auflösung ($< 0,1$ mm). Der Zeitraum (I) zeichnet sich durch ein verringertes p_0 , erhöhte Skalierungsparameter und Autokorrelationswerte, große Abweichungen in der empirischen Verteilungsfunktion geringer Intensitäten sowie veränderte ereignisbezogene Statistiken aus. Weiterhin finden sich Hinweise, dass selbst die Extremwerte einer hohen zeitlichen Aggregation (5 min) beeinflusst sein könnten.

Die Zielsetzung von NiedSim ist die Generierung einer realistischen und robusten synthetischen Zeitreihe für Kanalnetzsimulationen. Je nach Anwendungsgebiet sind unterschiedliche Eigenschaften der Niederschlagsleistungsdaten relevant. Für eine hydraulische Kanalnetzdimensionierung sind große Niederschlagsintensitäten von Bedeutung, wohingegen für Fragestellungen in der Schmutzfrachtsimulation die mittleren, langjährigen Eigenschaften eine entscheidendere Rolle spielen.

Eine Möglichkeit, synthetische Zeitreihen zu validieren, ist eine indirekte Validierung über Kanalsimulationen. Das hierfür verwendete, fiktive Kanalnetz basiert auf einem veränderten und erweiterten Modellansatz von *Drechsel* (1991) und spiegelt realistische Kanalnetze nach *ATV-A128* (1992) wider. Hierbei werden nicht nur der spezifische Regenabfluss und das spezifische Volumen von Überlaufbauwerken (mit und ohne Speicher) verändert, sondern auch die Länge des Oberflächenabflusses. Das fiktive Kanalnetz wird in dem Modell *Storm Water Management Model* (SWMM) implementiert, wobei der Oberflächenabfluss mit der Kontinuitäts- und Gauckler-Manning-Strickler-Gleichung berechnet und der Kanalabfluss hydrologisch simuliert wird. Das Kanalnetz wird hauptsächlich in zwei Varianten angewandt, die sich durch unterschiedlich lange Oberflächenabflüsse unterscheiden. Die eine Variante repräsentiert schnell reagierende Kanalnetze, wohingegen die andere Modifikation Netze widerspiegelt, bei denen der Abfluss stark gedämpft ist.

Eine Vergleich der mittleren jährlichen Entlastungsereignisse des Kanalnetzes Braunschweig mit dem fiktiven Kanalnetz weist darauf hin, dass beide Varianten notwendig sind, um unterschiedliche, reale Systeme abzubilden. Die Gegenüberstellung der absoluten Schmutzfrachtentlastungen mit einer statistischen Studie von *Brunner* (1997) zeigt, dass mit Hilfe der beide Varianten die vorhandenen Kanalnetze Baden-Württembergs gut abgebildet werden können.

Allgemeine Schlussfolgerungen zwischen Niederschlagsstatistiken und Mischwasserentlastungen

Das fiktive Kanalnetz wird zunächst umfassend mit Niederschlagsmessdaten untersucht, um zum einen deren zeitliche und räumliche Unsicherheiten zu ermitteln und zum anderen Konfidenzintervalle aufzustellen, die als Richtwerte für synthetischen Zeitreihen dienen. Die zeitliche Stichprobenunsicherheit wird mit einer Bootstrap-Methode berechnet. Die Ergebnisse unterstreichen die Probleme mit kurzen Zeitreihen (< 10 Jahre), die in der Literatur aufgezeigt werden. Allerdings können noch immer Unsicherheiten von bis zu 10 % auftreten, selbst wenn die Zeitreihen deutlich länger sind (37 Jahre).

Weiterhin wird in der Studie mit Hilfe einer Q-Q-Transformation gezeigt, dass die hohe zeitliche Stichprobenunsicherheit eher durch eine große jährliche Variabilität in den stündlichen Randverteilungen bedingt ist und weniger durch eine variable zeitliche Struktur. Der Einfluss der stündlichen Randverteilung auf Entlastungsereignisse ist wichtiger als die zeitliche Abfolge der Werte, allerdings nur, wenn die Struktur realistisch abgebildet ist. Andernfalls können erhebliche Abweichungen auftreten, selbst wenn die Randverteilungen identisch sind.

Wenn die Randverteilung eines Standortes mit Hilfe einer Q-Q-Transformation auf die Struktur eines anderen Standorts übertragen werden, zeigen sich große Abweichungen in den Entlastungsereignissen. Die Abweichung der stündlichen zeitlichen Struktur verschiedener Standorte ist größer als die jeweiligen Unterschiede zwischen den Jahren. Allerdings ist die räumliche Variabilität geringer als die zeitliche Stichprobenunsicherheit eines Standorts, wenn die Randverteilung mitbetrachtet und kurze Zeitreihen verwendet werden.

Die Struktur der Werte innerhalb einer Stunde ist insbesondere für Überlaufbauwerke ohne Speicherfunktion bei Kanalnetzen mit kurzen Reaktionszeiten von Bedeutung. Andernfalls führen Speicherungs- und Dämpfungsprozesse zu einem gleichmäßigeren Abfluss und egalisieren die stündliche zeitliche Struktur der Eingangsdaten.

In vielen Fällen ist der Einfluss der partiellen Serien, das heißt der Extremereignisse, auf die Schmutzfrachtsimulation weniger relevant als andere Ereignisse, die nicht über IDF-Kurven abgebildet werden können. Insbesondere beeinflussen partielle Serien die Entlastungsereignisse nur dann, wenn lange Dauern (> 1 h) betrachtet werden. Hingegen haben große Intensitäten innerhalb einer Stunde, bedingt durch die geringe kumulative Niederschlagssumme, kaum einen Einfluss auf Schmutzfrachtentlastungen.

Niederschlagsdaten sind sehr schief verteilt. Beispielweise sind nur ungefähr 20 % aller stündlichen Werte größer als 1,0 mm, enthalten jedoch 70 % der gesamten Niederschlagssumme. Trotzdem können kleine stündliche Werte ($< 1,0$ mm) oder fünfminütige Werte ($< 0,2$ mm, abhängig vom Messgerät) noch einen großen Einfluss auf die Entlastungsereignisse haben. Dies liegt daran, dass kleine Niederschlagswerte zum Abfluss beitragen können, wenn sie Teil eines größeren Niederschlagsereignisses sind. Deswegen sollten höhere Grenzwerte der Niederschlagsintensität nur angewandt werden, wenn ereignisbezogene Statistiken verwendet oder hohe Aggregationen betrachtet werden (z.B. Tageswerte). Für globale Statistiken in hoher Auflösung und für NiedSim, sind kleine Niederschlagswerte bedeutsam. Ein Grenzwert muss deswegen mit Bedacht gewählt werden.

Das Entlastungsvolumen ist in der Praxis der wichtigste Parameter bei den Berechnungen von Mischwassersystemen. Wie in der Studie gezeigt wird, ist dieser Parameter robuster gegenüber den Niederschlagseingangsdaten im Vergleich zur Entlastungsdauer. Da die Dauer auch von kleineren Werten beeinflusst wird, kann eine Validierung mit der Entlastungsdauer weitere Defizite in der zeitlichen Struktur der Niederschlagszeitreihen aufzeigen. Entsprechende Unstimmigkeiten sind nicht nur relevant, wenn die Entlastungsdauer direkt simuliert werden soll, sondern können auch für weitere Anwendungsgebiete wichtig sein, die nicht explizit in dieser Studie validiert werden.

Abschließend wird ein neues Maß entwickelt, die *asymmetry of dependence*, womit gezeigt wird, dass Niederschlagszeitreihen irreversibel sind. Diese Irreversibilität beeinflusst die Abflusseigenschaften und somit die Überlaufstatistiken von Schmutzfrachtsimulationen. Der Einfluss liegt in derselben Größenordnung wie die Abweichungen der synthetischen Zeitreihen von NiedSim von gemessenen Reihen.

Güte der stündlichen Optimierung von NiedSim

Die Güte der stündlichen Optimierung von NiedSim wird durch einen direkten Vergleich der simulierten Zeitreihe mit den Zielstatistiken beurteilt. Weiterhin werden die Daten mit Hilfe von partiellen Serien, dem fiktiven Kanalnetz und ereignisbezogenen Statistiken indirekt validiert. Um den Einfluss der stündlichen Randverteilung zu vermeiden, wird nur die zeitliche Struktur einer Zeitreihe mit gemessenen, jedoch zufällig angeordneten Werten, optimiert, ohne die Randverteilung explizit zu simulieren. Die Reihen zeigen große, systematische Abweichungen in der indirekten Validierung, obwohl die Zielstatistiken in den meisten Fällen gut optimiert werden können. Die Ergebnisse unterstreichen die Komplexität der Niederschlag-Abfluss-Überlauf-Prozesse, welche weder mit wenigen Niederschlagsstatistiken, noch mit der alleinigen Betrachtung der partiellen Serien oder durch die Simulation von einigen wenigen Überlaufbauwerken abgebildet werden können.

Die Optimierung der Autokorrelation verursacht eine starke Anhäufung großer Niederschlagswerte. Aus diesem Grund wird ein Poisson-Prozess eingeführt, welcher große stündliche Niederschläge vor der Optimierung fest in der Zeitreihe positioniert. Die Anzahl an gesetzten Werten pro Monate hängt von der Jahreszeit und der monatlichen Niederschlagssumme ab. Es werden jedoch nicht alle großen Werte fest positioniert, um das Auftreten von mehreren großen Werten innerhalb eines Ereignisses zu ermöglichen. Die Parameter des Poisson-Prozesses sind bezüglich ihrer Auswirkung auf die Entlastungsereignisse relativ robust. Außerdem sind sie unabhängig von Orts- und Höhenlage, weswegen ein einziger Parametersatz für Süddeutschland angepasst wird. Die Einführung des Poisson-Prozesses reduziert nicht nur die Abweichung in der Zeitreihe für große Werte, die durch partielle Serien dargestellt werden, sondern verbessert auch die vollständige zeitliche Struktur, was sich durch verringerte Abweichungen in den Entlastungsereignissen zeigt.

Zusätzliche Probleme können mit einer fehlerhaften Abbildung der Intermittenz in Verbindung gebracht werden. Es werden zu viele kurze Ereignisse mit großen Niederschlagsvolumina generiert, wohingegen die Niederschlagsmenge von langen Ereignissen deutlich unterschätzt werden. Aus diesem Grund wird die Häufigkeitsverteilung der Niederschlagsereignisse als zusätzlicher Parameter in die Zielfunktion mit aufgenommen. Die Validierung der Ereignisstatistiken zeigt eine Verbesserung der Ereignisstruktur, obwohl der positive Effekt auf die Entlastungsereignisse nicht an jedem Ort ausgeprägt ist.

Mit der Einführung des Poisson-Prozesses und der Häufigkeitsverteilung der Niederschlagsereignisse kann die zeitliche Struktur der Zeitreihen im Vergleich zur früheren Version von NiedSim verbessert werden. Die noch vorhandenen Abweichungen in der partiellen Serie und im Entlastungsvolumen sind im Vergleich zur Stichprobenunsicherheit klein. Allerdings kann die Entlastungsdauer je nach verwendetem Kanalnetz noch immer systematisch unterschätzt werden. Ein Zusammenhang dieser Abweichungen mit weiteren Statistiken wie Verteilungsfunktionen oder bivariaten, ereignisbezogenen Statistiken kann jedoch nicht gezeigt werden.

Güte der fünfminütigen Optimierung von NiedSim

In der Studie werden mehrere Herausforderungen bei der Disaggregation von stündlichen Niederschlagsdaten aufgezeigt. Zunächst kann der Einfluss von unterschiedlichen Messgeräten auf die Statistiken von hochaufgelösten Daten nicht vernachlässigt werden. Aus diesem Grund sollten nur Daten der neusten Messgeräte (z.B. digitale Niederschlagswaagen) verwendet werden, um homogene Statistiken zu erhalten. Die Stichprobenunsicherheit solcher Daten ist an vielen Orten auf Grund der reduzierten Länge verfügbarer Messdaten erhöht. Dies könnte für die Berechnung der mittleren jährlichen Eingangstatistiken in NiedSim vernachlässigt werden, da diese relativ robust sind. Allerdings beeinflusst diese Unsicherheit sowohl die Langzeit-Validierung mit Kanalnetzmodellen als auch die partiellen Serien.

Eine weitere Schwierigkeit besteht in der Rekonstruktion der hochaufgelösten zeitlichen Struktur. Bereits die Güte der individuellen fünfminütigen Optimierung der Autokorrelation oder der Skalierungsparameter erweist sich als schlechter im Vergleich zur stündlichen Optimierung. Sobald beide Parameter gleichzeitig optimiert werden, verringert sich die Güte weiter und hängt von der gewählten Gewichtung in der Zielfunktion ab. Die besten Ergebnisse werden mit einer Optimierung von Skalierungsparametern und gewichteter Autokorrelation erreicht, wobei der Fokus auf der Korrelation liegen sollte.

Trotzdem sind systematische Probleme in der zeitlichen Struktur vorhanden. Das heißt, die zu wenigen Null-Werte auf höheren Auflösungen deuten darauf hin, dass die stündlichen Niederschlagshöhen jeweils auf alle fünf Minutenwerte verteilt werden, anstatt Ereignisse mit einem eindeutigen Beginn und Ende zu generieren. Eine zusätzliche Optimierung auf die Verteilung der stündlichen Entropie-Werte kann die fünfminütige Struktur nicht verbessern und zeigt grundsätzliche Probleme in der Optimierung fünfminütiger Zeitreihen mit Hilfe von NiedSim auf. Allerdings können mit der NiedSim Disaggregation trotzdem Verbesserungen in der Zeitreihe im Vergleich zu einer Reihe mit zufällig veränderten fünf Minutenwerten erzielt werden. Außerdem zeigt die Studie, dass die partielle Serie nachträglich angepasst werden kann, ohne die Entlastungsereignisse wesentlich zu beeinflussen.

Abschließend kann zu folgendem Ergebnis gekommen werden. Die Schwierigkeiten in der Disaggregation sind für viele Mischwassersysteme im Vergleich zur relativ gut funktionierenden stündlichen Generierung von untergeordneter Bedeutung. Da außerdem die großen Niederschlagsintensitäten der partiellen Serie nachträglich angepasst werden könnten, ist davon auszugehen, dass die Optimierung der zeitlichen Struktur mit Hilfe von NiedSim für die meisten Kanalnetz Anwendungen geeignet ist.

1. Introduction

Human settlements have traditionally been founded close to rivers or at other places where enough water was available for drinking water, irrigation of plants and watering animals. The waters have also been used to dispose waste and waste water for centuries. On the other hand, water can be a source of potential hazard causing floods and spreading diseases. Therefore, the discharge of water has been a challenge since the first settlements and triggered the concept of constructing sewers for the collection of waste water as well as for draining the settlement from rainfall.

Great progress was achieved during the 19th and 20th century in urban sanitation in the industrialized world leading to the development of modern sewer systems. Back in the days when no computers existed or when their performance was too weak, the sewers were constructed based on simple manual calculations of the expected discharge. Nowadays, the target in sanitary environmental engineering changes more and more from simple, fast drainage of storm and waste water to integrated concepts including the protection of the environment. Due to the high complexity of rainfall-runoff processes, such integrated concepts can only be realized using sewer system simulations. They are not only needed for integrated approaches, but also for the planning of new or the revision of existing sewers. In particular, the latter one gets increasing attention because around 10 % of all existing sewers in Germany were built more than 100 years ago (*Berger and Falk, 2009*). Sewer systems are an expensive part of urban infrastructure with a replacement value of EUR 687 billion in Germany (*Berger and Falk, 2009*). Consequently, large investments must be spent on sewers for maintenance. Because their planning horizons are several decades, the risk of an expensive bad investment can be high. By using sewer simulations the sewers can be planned properly and designed efficiently resulting in a better flood protection and waste water discharge, in addition with reduced ecological impacts and less investment costs.

One major input of sewer models are the characteristics of local precipitation. They can be reflected best with observed, long-term rainfall time series of high data quality. Unfortunately, such time series are not available for many locations. A solution can be synthetic precipitation time series that replace the missing observations. A precipitation simulator developed at the University of Stuttgart is the *NiederschlagsSimulator* (NiedSim) [precipitation simulator], which has been in operational use since the year 2000. NiedSim generates high resolution time series stochastically based on measured statistics. During the development of NiedSim, the model was validated using a variety of different statistics including extreme value characteristics which could be reproduced well.

In a preliminary study (*Müller et al., 2014*), I applied synthetic time series of NiedSim to a combined sewer model. I illustrated large deviations of the CSO volume of synthetic time

series of NiedSim compared to measured ones and found the temporal structure to be a main challenge.

Another concern with the current version of NiedSim arose during the data update in 2013. As the operational version of NiedSim was developed at the beginning of the years 2000, around ten years of additional high resolution data were available, but had not been used until then. I found that the performance of the algorithm varied depending on the measured time periods of the observed data because the calculated input statistics are affected by different measurement devices.

Both results were the motivation of my present study to reexamine the current version of NiedSim. The main objective is to enhance the NiedSim algorithm towards a more realistic temporal structure of the synthetic precipitation time series such that they can be applied for sewer simulations. For this purpose, I performed a comprehensive analysis of measured precipitation to determine the important characteristics for sewer simulations. Furthermore, the examination of measured data can show potential error sources in the input data, can illustrate influences of different measurement devices and finally, can help to define natural limits of accuracy. Therefore, beside the development of the NiedSim algorithm, the focus of my study is on an extensive analysis of measured precipitation data with respect to rainfall modeling and sewer simulations.

My study is structured as follows. **Chapter 2** gives a general overview of precipitation including rainfall measurement devices and its statistical description. The observed rainfall data used in my study is analyzed statistically in **Chapter 3**. In **Chapter 4**, I introduce the general concepts of the simulation of synthetic precipitation as well as the theoretical background of the current version of NiedSim. I discuss the applications of synthetic rainfall in urban hydrology in **Chapter 5**. For the validation of synthetic rainfall, I apply an artificial sewer system which I explain together with the used sewer software in **Chapter 6**. Before I make use of the artificial sewer to validate synthetic data, I perform a comprehensive study of the sewer using measured precipitation data in **Chapter 7**. The following two chapters concern the actual developments of the NiedSim algorithm as well as their respective validation. In **Chapter 8** I investigate the optimization of the hourly structure, whereas in **Chapter 9** I analyze the disaggregation of hourly time series to series of 5 min resolution. In **Chapter 10** I discuss an additional property of precipitation time series, the irreversibility, and show its impact on combined sewers. In the last **Chapter 11**, I finally discuss the results and propose further investigations in an outlook.

2. Observed precipitation

The major input variable of model applications in the field of urban hydrology and urban drainage is precipitation. Precipitation is a collective term that combines the water equivalent of rain, hail, snow, dew, rime, hoar frost and fog precipitation (Jarraud, 2008). The amount of precipitation is defined as the water depth that covers an impervious, flat area if evaporation and runoff is neglected. That is, 1 liter of water volume falling on an area of 1 m^2 results in 1 mm of water depth. The precipitation amount is usually referred on the time interval in which the precipitation occurred and is then expressed as an intensity. The precipitation intensity is defined as precipitation depth per unit time interval (Lanza *et al.*, 2005). One should differentiate between *instantaneous rainfall rate*, which is usually referred to as rainfall intensity, and *mean rainfall rate* (Dunkerley, 2008a). The later can be expressed as an instantaneous rate that is uniformly extrapolated to a duration longer than the actual instantaneous time step (usually mm/h). In the context of precipitation events, the mean rainfall rate is defined as the integral depth over the event divided by the event duration (Dunkerley, 2008a). In my study, I refer the intensity to the *instantaneous rainfall rate* if not explicitly stated otherwise. Furthermore, I do not distinguish between rainfall and (solid) precipitation, and I use both words as synonyms.

2.1. Overview of measurement devices

The conventional way of measuring precipitation are point measurements using rain gauges, where the precipitation is collected and measured over a well defined area (e.g. in Germany 200 cm^2). The most simple rain gauge is the manual gauge, where the water is collected in a container. The container is emptied once per day and the accumulated amount of water is quantified. The quantification of daily precipitation is rather exact, however, the distribution within a day, that is the temporal structure, is not measured.

In the nineteenth century, so called pluviometers have been developed that allow to quantify the intra-daily intensity. The first pluviometers were based on paper charts records (Strangeways, 2006). They record the amount of rainfall with a moving pen that marks the amount of rainfall on a chart. The movement is induced either by a floating system (Hellmann) or by a weighing system. In the Hellmann measurement, the raising water level in the collection container causes a floating pen to move. In the weighing system, the pen arm is attached to a container which is positioned on a spring. The increasing amount of water in the container during rainfall forces the pen to change its position. Both system are continuous measurements. However, as the paper records are analogue measurements, the charts have to be digitized afterwards for further usage.

Since the development of modern computers, digital data loggers have been available. They are used instead of paper charts and allow fully automated rainfall measurements. Digital measurement systems can be subdivided into three main types: tipping-buckets, weighing systems and drop-counters (*Strangeways, 2006*).

The most widely used type is the tipping-bucket. It is a non-continuous measurement that consists of two buckets with a predefined size. When it rains, only one side can be filled with water. As soon as the maximum volume is reached, the bucket tips, and the second bucket can be filled until it tips again back to the original position. A tip is designed such that it represents a certain amount of precipitation. The number of tipplings within a time period can then be referred to the intensity of precipitation. For the tipping bucket the maximum resolution is usually 0.1 mm. The weighing system works similar to the analogue case, except that the weight is directly converted into an electrical signal. In drop-counting measurements, the precipitation is first transformed into single drops of a predefined size, which are then counted by optical sensors. The resolution of weighing and drop-counting pluviometers may be 0.01 mm or even higher.

The *Deutscher Wetterdienst* [German Meteorological Service] (DWD) installed the first generation of automated pluviometers in the early 1990s. In the year 2002, the DWD started a successive upgrade of their pluviometers to the second generation of automated pluviometers, the so called AMDA pluviometers (Automatische Messdatenerfassungsanlage) (*Behrendt et al., 2011*).

All measurements described above are point measurements. Area-related precipitation data can be obtained by radar, satellite or microwave links. These measurements are not further discussed here because this work is only concerned with single time series based on data from pluviometers.

2.2. Statistical characterization of precipitation

Precipitation is the result of very complex and highly non-linear atmospheric processes. Most precipitation is either caused by frontal systems or by convective storms (and in mountainous areas additionally by orographic lifts). Frontal systems usually come along with stratiform rainfall of moderate intensity lasting days, whereas convective storms can occur suddenly with heavy intensities, but remain only for a few hours. Both processes, however, often occur not clearly separated from each other, but as a complex result of mutual interactions. Although both processes have different genesis, and thus different precipitation characteristics, the resulting measurable output, the precipitation, is the same. Therefore, the characteristics of a precipitation time series depend on the regarded temporal aggregation and is difficult (if not impossible) to be fully described by a few parameters.

On the one hand, the time dependent process is rather persistent depending on the temporal aggregation. In Germany, for example, if it rains at time step t at a five minute aggregation, it is very likely that it will also rain at the next time step $t + 1$. Even at a daily aggregation, the probability of rain at the day after a wet day is increased indicated by larger Markov transition probabilities (*Bárdossy, 1993*).

Global statistics	Event based statistics	Extreme value statistics
(ϕ) max., mean, var., skew. (different scales)	wet spell -depth, -duration -intensity, -frequency internal peak, time to peak	intensity-duration-frequency (IDF) maximum / partial
(ϕ) distribution functions (diff. scales)	dry spell -duration, -frequency	depth-duration-frequency (DDF) maximum / partial
(ϕ) p_0 (probability zero prec.) (diff. scales)	antecedent precipitation	
(ϕ) entropy (diff. scales)		
(ψ) autocorrelation functions (diff. scales / time lags)		
(ψ) scaling parameters (diff. moments)		

Table 2.1.: Overview of different precipitation statistics.

On the other hand, the rather stable rainfall process may exhibit abrupt changes. For example, if a small convective storm cell occurs, a dry day may be interrupted by one or two hours of very intense precipitation. The combination of a stable process that is interrupted by sudden changes is called intermittent process. The intermittency is an important characteristic of precipitation, but its proper description is very challenging.

Due to the very complex temporal structure of precipitation time series, no common rules or methods have been defined to characterize precipitation in the literature so far. Instead, many different statistics have been developed and applied in the past. In Table 2.1 I summarize the most commonly used statistics and group them into three categories: global, event-based and extreme value statistics.

2.2.1. Global statistics

Global statistics are calculated by using the entire time series at a specific temporal aggregation. I divided them further into two subgroups indicated with ϕ and ψ . The first group ϕ are statistics describing the marginals of precipitation values at a certain time scale. They can be single values like the maximum, or statistics based on moments like the mean, the variance (standard deviation) or the skewness of time series. Alternatively, empirical distribution functions (EDFs) can be calculated which reflect the non-exceedance probability (quantile) of precipitation values. Another often applied statistic is the probability of a time step having zero precipitation p_0 . If only one aggregation is considered, none of the measures ϕ include any information about the temporal order of the values. Only if several time scales are regarded, one can refer indirectly on the temporal structure. For example, if p_0 is evaluated at several different scales, it can be used to describe the intermittency (e.g. *Molnar and Burlando, 2005*). The second group ψ contains all parameters describing the temporal structure directly. The most important ones are the autocorrelation functions (ACFs), the scaling parameters and the entropy.

Autocorrelation The autocorrelation $r_l^{(j)}$ describes the correlation of consecutive values in a time series separated by a time lag (Wilks, 2011), and is calculated as

$$r_l^{(j)} = \frac{\sum_{i=1}^{T-l} (x_i^{(j)} - \bar{x}^{(j)})(x_{i+l}^{(j)} - \bar{x}^{(j)})}{\sum_{i=1}^T (x_i^{(j)} - \bar{x}^{(j)})^2} \quad (2.1)$$

where x_i is a precipitation value at time step i , \bar{x} is the mean precipitation value, l is the lag, j is the aggregation and T the length of the time series. The lag defines the shift of the time series and reflects the regarded length of the temporal memory. The autocorrelation can be interpreted as the temporal stability of a time series, that is, the higher $r_l^{(j)}$ is, the more similar two successive values are. The autocorrelation can be highly influenced by large values as I show later in the study.

Scaling parameters Another way of analyzing the temporal structure is by calculating statistical moments at different scales and analyzing their dependence across the scales. The scaling parameters do not show directly the temporal structure at a specific aggregation, but relate the changing moments at different scales, and thus represent the temporal structure indirectly via the aggregation process.

The k -th empirical moment m for an aggregation j can be calculated as (e.g. Lombardo et al., 2014)

$$m_k^{(j)} = \frac{1}{T} \sum_{i=1}^T (H_{wet,i}^{(j)})^k, \quad (2.2)$$

where $H_{wet,i}^{(j)}$ is a precipitation value larger than zero at time step i and T is the length of the wet time series. The relationship between the k -th moment and the aggregation level j is assumed to follow a power law which can be described by

$$m_k^{(j)} = a_k * S_j^{b_k}, \quad (2.3)$$

where S_j is the j -th aggregation level in minutes and a_k and b_k are the parameters of the power function. A linear regression of the moments and the aggregation level can be applied in the log-space. The parameter a_k is a simple constant, whereas the actual dependence between aggregation and moments is described by the parameter b_k . In the following, the parameter b_k is referred to as scaling parameter. If the scaling parameters of different moments show a linear relationship, the process is known as simple scaling, otherwise as multiple scaling (Burlando and Rosso, 1996).

Entropy The entropy is a statistic which can be used to describe the information content of a random variable, its probability of occurrence and its uncertainty (*Singh, 2011*). All characteristics are related to each other. The more likely a certain realization of a random variable is, the less uncertain it is, and thus the less additional information can be provided by other realizations. A low entropy characterizes such a random variable. On the other hand, the more diverse realizations can occur, the lower their individual probabilities are, which leads to a higher uncertainty of the individual realizations. Therefore, the information content is increased and is described by a high entropy. The general form of the entropy H is defined as

$$H(X) = - \sum_{i=1}^n p(x_i) * \log(p(x_i)) \quad (2.4)$$

where $p(x_i)$ represents the probabilities of a realization x of the random variable X , and n is the sample size (*Singh, 2011*).

The general concept of entropy can be applied to precipitation. Here, $p(x_i)$ does not describe a probability, but defines a relative value describing the information content at a high temporal resolution k (e.g. 5 min aggregation) with respect to a coarser resolution j (e.g. hourly aggregation). The more the information (precipitation) is spread across single increments, the higher the information content and, thus, the entropy is. For this purpose, the Shanon entropy $H_S(x)^{(j)}$ (defined by using \log_2) of a precipitation value x at an aggregation j is calculated using the relative contributions $p_i^{(k)}$ at the finer aggregation k . Mathematically, this can be expressed as

$$H_S(x)^{(j)} = - \sum_{i=1}^n p_i^{(k)} * \log_2(p_i^{(k)}) \quad (2.5)$$

with

$$p_i^{(k)} = \frac{x_i^{(k)}}{x^{(j)}} = \frac{x_i^{(k)}}{\sum_{i=1}^n x_i^{(k)}} \quad (2.6)$$

where x represents the precipitation values at coarse (j) and fine aggregation (k), and n defines the number of fractions in which the coarse aggregation is subdivided. Here, the binary logarithm is used to calculate the Shanon entropy.

The Shanon entropy has values on the interval $[0 \leq H_S^{(j)} \leq \log_2(n)]$. For example, if the hourly entropy is calculated using 5 min data, $H_S^{(H)}$ is defined on $[0 \leq H_S^{(H)} \leq 3.58]$. That is, if just one single 5 min value occurs and all other values within the hour are zero, the 5 min value is equal to the hourly sum. Consequently, the respective hourly entropy is zero implying least uncertainty because all information is provided by one value. In the other case, where the hourly rainfall is distributed uniformly within 5 min values, each 5 min value contains some information, and thus the entropy is at its maximum, here 3.58. Therefore, the entropy at a certain time scale contains information about the precipitation distribution at a sub-scale.

2.2.2. Event-based statistics

Event based statistics do not describe the time series as a whole, but characterize well-defined parts, the precipitation events, with statistics like depth, duration, frequency, mean and peak intensity. Furthermore, the time between two events, the dry spell duration, can be obtained as well as information about the antecedent precipitation. If two events are separated they are assumed to be independent. The main challenge in this context is the definition of the event separation criteria. If the interim time between two events is short, even individual convective thunderstorms can be separated and analyzed. However, by applying a short interim time, long lasting events will also be split, although they might be part of the same large scale synoptic weather phenomena and, thus, should not be considered as independent. On the other hand, if the interim time is too long, several short events will be combined to one long event. The result is a decreased variability in the event statistics due to the loss of information about the internal intensity structure of the event. The definition of the interim time is, hence, a trade-off between event independence and representation of the intensity characteristics.

Many different event separation methods exist. In almost all cases not the raw time series is used, but a time series with thresholds for minimum precipitation intensity and/or depth (e.g. *Xanthopoulos, 1990*). Otherwise, even very small precipitation values would be considered as individual events. The major independent criterion is the interim time describing the time between two individual events. The variable can be defined using pure statistics or by a subsequent application. For example, *Restrepo-Posada and Eagleson (1982)* use the coefficient of variation, which is the ratio of the standard deviation and the mean, to define the interim time between independent storms. Whereas *DWA-A118 (2006)* proposes a time that should depend on the discharge time of the drainage system, but at least four hours. *Dunkerley (2008b)* shows that different event separation criteria result in varying event statistics. Consequently, if not the same criteria are applied, one cannot compare event-based statistic of different studies. The missing uniqueness and the lacking general definition are major drawbacks of event-based statistics.

In my study, I distinguish between two definitions: The advanced definition of a proper precipitation event is indicated as *event*, whereas a simple separation of consecutive wet values by at least n dry time steps is defined as *spell*.

2.2.3. Extreme value statistics

The distribution of precipitation values exhibit a high positive skewness at all aggregations from 5 min to 24 h, that is, the larger the values the less frequent they occur. Nevertheless, the rare precipitation events cannot be neglected. Indeed, they are very important because of the disastrous consequences for urban infrastructure in the case of their occurrence. Therefore, the extreme value statistics, describing the upper tail of the distribution function, is an essential part of the statistical analysis of precipitation. However, one should keep in mind that extreme value statistics by definition characterize only very few precipitation values. For example, the average number of values per year usually used in extreme value statistics

is about 3 out of 105,000 at the 5 min aggregation. Even if only wet values are considered (90 – 95 % of the 5 min values are zero), the extreme value statistics reflect less than 0.05 % of all wet 5 min values. For larger aggregations the number of 5 min values contributing to extremes increase, but still relatively few values are taken into account ($\approx 1\%$ at a daily aggregation).

The extremes are usually evaluated by a relationship between intensity, duration and return period and are illustrated using intensity-duration-frequency (IDF)-curves. Less often depth-duration-frequency (DDF) curves are used for the same purpose. In such curves, the precipitation intensity (or depth) is plotted against its empirical return period (= frequency) for a given aggregation (= duration).

The DWD provides an extreme value analysis of entire Germany (KOSTRA-DWD-2000) which is described in *Bartels et al.* (2005) and based on *DWA-A531* (2012). The IDF relationships in KOSTRA are evaluated for durations from 5 min to 72 h and return periods of 1 year to 100 years on a 8.45×8.45 km raster. Although the DWD recently released an update, KOSTRA-DWD-2010 (*Malitz and Ertel*, 2015), the term KOSTRA refers to KOSTRA-DWD-2000 in my study if not stated otherwise.

The precipitation intensities can be obtained by using the maximum of each year (annual series), or by taking a number of highest values within the entire time period (partial series). In both methods, the intensity series are obtained by considering the largest intensities calculated with a moving time window of the desired duration. A major assumption of the extreme value analysis is the independence of these values. The independence can be assumed to be automatically fulfilled in the case of annual maxima. If the partial series is used, several approaches exist to assure the independence of values being located in the same year. *Verhoest et al.* (1997) defines values to be independent if at least 24 dry hours separate the extremes. Whereas in *DWA-A531* (2012) the criteria is weaker, that is, the interim time must only be the length of the regarded aggregation, but at least four hours.

For the partial series usually a peak over threshold (POT) approach is applied, where all values above a given threshold are used for the extreme value analysis. The POT threshold can also be set indirectly by explicitly defining the number of largest values. For example, in KOSTRA the Euler number $e \approx 2.718$ of values per year are used to define the partial series. According to *Verhoest et al.* (1997), the partial series are more reliable for return periods of 0.1 to 10 years compared to the annual series. As these return periods are most important for urban hydrology, I use the partial series in this study instead of the annual series.

The return period T for a given intensity I and duration D is defined as the time interval in which the intensity $I(D)$ occurs on average at least once. The return period T can be expressed mathematically as the reciprocal of the exceedance frequency or in terms of non-exceedance frequency F : $T = 1/(1 - F)$. That is, the return period T is directly related to the EDF of the intensities. To obtain an empirical IDF-curve, the ranked intensities of the maximum or partial series are plotted against their respective empirical frequencies (return periods). Plenty of different methods exist in literature to obtain the plotting positions for a given sample size L (e.g. Hazen or Gumbel plotting positions in *Cunnane* (1978)). However, according to *Makkonen* (2006) only the Weibull plotting position, $T = (L + 1)/k$,

represents the extreme values correctly independent of the underlying distribution, where k is the index of the sample values in descending order. Nevertheless, in order to be comparable to KOSTRA based on *DWA-A531* (2012), I use the formula defined there, that is, $T = (L + 0.2)/(k - 0.4) \cdot M/L$, where M is the length of the time series in years.

In a final step, a theoretical distribution function can be fit to the empirical IDF relationship between $I(D)$ and T . A distribution of the extreme value distribution families (type I, II and III) can be used, for example the exponential, the generalized extreme value (GEV), the Gumbel or the Gamma distribution (*Koutsoyiannis et al.*, 1998). In the KOSTRA analysis the exponential distribution is applied. When the IDF-curves are fitted individually for each duration as applied in KOSTRA, they will probably not be consistent across all durations. Therefore, a parameter adjustment is necessary as explained in *DWA-A531* (2012) in more details. In my study I do not fit a theoretical distribution, but compare only the empirical IDF-curves.

2.2.4. Other approaches

In addition to the above mentioned statistics, several other methods have been developed to describe precipitation. One of the first methods were the mass curves developed by *Huff* (1967) that are still used (e.g. *Dolšak et al.* (2016)). *Huff* (1967) grouped precipitation events into four quarters depending on the occurrence of the maximum rainfall intensity. Each quarter can then be characterized by the cumulative distribution of precipitation amounts with its respective cumulative distribution of durations.

Another possibility to describe the temporal structure is the spectral analysis of Fourier-transformed time series (*Burgueño et al.*, 1990; *Paschalis et al.*, 2014) as the spectral density is related to the autocorrelation function. The power spectrum, which is the spectral density as a function of the frequencies, can be plotted and is typically evaluated graphically (*Wilks*, 2011).

A way to describe the intermittency across a large range of scales was recently developed by *Schleiss and Smith* (2016). The intermittency is characterized based on inter-amount times. The dispersion of the inter-amount times is reflected by a burstiness coefficient, and the temporal ordering is specified by a memory parameter.

More recently the evaluation using bivariate or multivariate copulas have been developed. They can be widely used to analyze the dependence structure of rainfall independently of their marginal distributions. The marginals are separated from the dependence structure by transforming the precipitation values into the rank space. Afterwards, all statistics can be evaluated in their rank space, like event-based statistics (*Salvadori and Michele*, 2006) or extreme value analysis (*Vandenberghe et al.*, 2011).

3. Analysis of observed precipitation

The focus of my study lies on the simulation of single point rainfall, and consequently precipitation of several stations are used as reference data individually. However, measured precipitation time series cannot be treated as a perfect representation of natural rainfall for two reasons. First, precipitation data is difficult to measure and a variety of measurement errors may occur. Second, several sources of uncertainty exist such that a measured time series must be regarded as one individual sample that does not fully represent the truth, respectively, the population of the underlying process. Therefore, the measured data should be extensively analyzed in order to be able to observe, determine and quantify the uncertainty. The purpose of this chapter is, first, to show the statistics of the input data that are the basis for the simulation of synthetic series later in my study, and second, to illustrate potential issues related to measured data.

3.1. Measurement errors and inhomogeneity

The measurement of precipitation is vulnerable to a variety of potential errors. As they can have significant influences on applications, I briefly point out the different sources as well as their relevance for the generation of synthetic precipitation.

Some measurement issues have a natural, physical reason, for example, errors induced due to wind drift, wetting loss and evaporation (*Sevruk and Klemm, 1989*). Such errors lead to systematic underestimations of the measured precipitation. The issues are well known and many studies have been carried out, how such errors can be reduced in situ or corrected afterwards (e.g. *Sevruk and Hamon, 1984*).

Furthermore, measured time series may exhibit unnatural errors, for example, due to bad maintenance, to damages in the devices or to failures in the electronics. These errors show an erratic characteristic and are supposed to be observed, marked or corrected by the operator. Unless the rainfall values caused by such errors are very unrealistic, they are hardly detectable without any further information.

Another important issue is the uncertainty related to different measurements devices. *Nystuen et al. (1996)* illustrates that measured precipitation of different devices is generally highly correlated, but the accuracy varies depending on the rainfall intensity. If the uncertainty has a random nature, it is less problematic because on the long run they cancel each other out. Even if certain devices show a systematic bias, one could account for such an error as long as their device and its bias is known.

For all errors described above additional information needs to be available, that is, knowledge about the device, the measurement process or the post-processing. As such information is not accessible in my case, I use the measured data without any additional corrections except for the ones already performed by the operator. However, I have no specific information about the applied post-processing methods. Consequently, there is still uncertainty in the measured data with respect to error handling. Nevertheless, such kind of time series would be used for applications in practice and any uncertainty would usually be indirectly considered by security surcharges. Therefore, the applied data is not supposed to reflect the reality of local rainfall, but considered as a realistic time series practically applied in urban hydrology.

Another issue of measured data is the inhomogeneity that may play an import role for the simulation of rainfall. Inhomogeneity could arise from several sources like different measurement devices, changed locations or difference in error handling and post-processing. Inhomogeneous time series can show systematic differences in the statistics between certain sections of the time periods. Depending on the length of the inhomogeneous parts and on the influence on the statistics, they can have high impacts on the rainfall simulation.

Therefore, I focus on the inhomogeneities in the observed time series in this chapter, rather than focusing on sources and specific treatments of measurement errors.

3.2. Overview of applied measured data

In my study, I use data in a high temporal resolution of 5 min of pluviometers provided by the DWD. The time series start around 1960 and last to 2012. However, the actual available period is shorter than 53 years because of gaps with missing years. As the measurement devices of the DWD changed during the time period twice (see Chap. 2.1), the long time period is split into three different sub-periods which I analyzed individually throughout the study. The separation enables to investigate the effect of inhomogeneities of time series on rainfall simulations. However, one should keep in mind that the reduced time periods come along with an increased sampling uncertainty.

I defined the three periods as follows: period I (196x – 198x), period II (199x – early 200x) and period III (late 20xx). The exact start and end of the periods is chosen depending on the data availability of the pluviometer in each period. Three locations, Freiburg, Kempten and Weiden, are used because of their large data availability. All three stations provide long precipitation data (at least 35 years in total) in a high temporal resolution covering all three time periods. Additionally, I use the station Augsburg because of its long data set in time period III. Table 3.1 summarizes the applied time periods for the different locations.

The three main locations are representative for different climatic conditions occurring in Germany. Kempten is located in the Allgäu close to the Alps and represents a wet location with more than 1200 mm of precipitation per year. Weiden is a city in the Upper Palatinate of Bavaria that reflects a rather dry climate with around 700 mm of precipitation per year. In Freiburg, situated on the windward side of the Black Forest in Baden-Württemberg, the annual precipitation amount is around 900 mm.

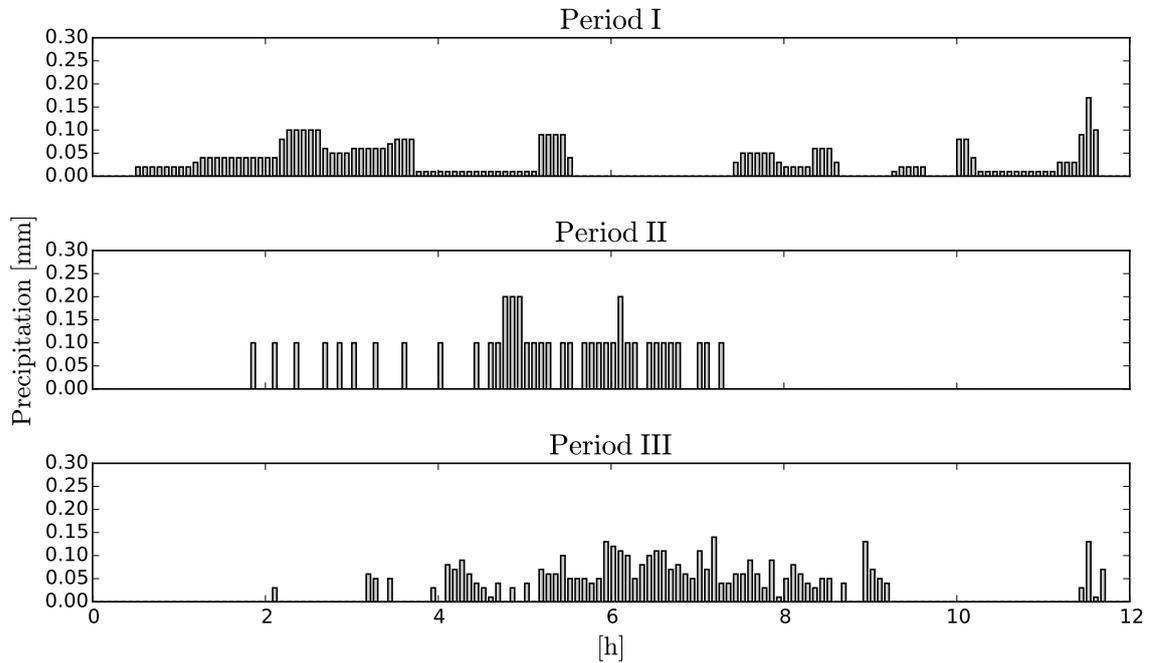


Figure 3.1.: 5 min time series of 12 hours for precipitation events of three periods. Each event sums up to around 4 mm of precipitation, but shows very different intensity characteristics at the 5 min aggregation caused by different measurement techniques.

Table 3.2 gives a basic overview of the most important statistics used in the study: the number of missing values n_{nan} , the mean annual precipitation sum ΣH , the number of wet values $n_{wet}^{(j)}$, the frequency of hourly wet spells with a length of one hour $ws_{1h}^{(j)}$, the frequency of all hourly wet spells $ws_{\Sigma}^{(h)}$, the probability of precipitation values being zero $p_0^{(j)}$ and the autocorrelation of lag 1 $r_1^{(j)}$. Additionally, the maximum, mean, standard deviation and skewness of the time series $H^{(j)}$ are shown. The superscript (j) denotes the temporal aggregation: five minutes (5min), one hour (h) or one day (d).

The data quality with respect to missing data is good. All locations exhibit a number of missing values n_{nan} of less than 1 % for period II and III. In period I, the number of missing values is significantly higher, where almost 50 % of all 5 min values are not available. However, if one considers the mean annual precipitation sum ΣH of period I, only a small systematic underestimation can be noticed compared to period II and III, or compared to

Location	Period I	Period II	Period III
Freiburg and Kempten	1961 – 1980	1995 – 2005	2007 – 2012
Weiden	1964 – 1981	1995 – 2005	2007 – 2012
Augsburg	-	-	1997 – 2012

Table 3.1.: Overview of the time periods used for different locations.

		F.: I	F.: II	F.: III	W.: I	W.: II	W.: III	K.: I	K.: II	K.: III	A.: III
n_{nan}	[%]	50.4	0.027	0.093	51.3	0.002	0.077	46.6	0.006	0.253	0.633
ΣH	[mm/year]	911	946	832	657	727	737	1242	1269	1186	758
ΣH_{daily}	[mm/year]	939	942	832	701	723	743	1253	1265	1191	761
$n_{wet}^{(5min)}$	[n/year]	9673	4843	6317	8445	4015	6690	14205	8960	9751	6288
$n_{wet}^{(h)}$	[n/year]	1067	1060	1106	979	1012	1218	1497	1523	1524	1119
$n_{wet}^{(d)}$	[n/year]	166	176	185	160	178	190	185	191	197	184
$ws_{1h}^{(h)}$	[n/year]	64	134	128	64	141	132	59	112	110	127
$ws_{2h}^{(h)}$	[n/year]	58	76	81	63	74	85	60	70	67	78
$ws_{\Sigma}^{(h)}$	[n/year]	250	338	344	248	342	364	279	352	344	338
$p_0^{(5min)}$	[-]	0.91	0.95	0.94	0.92	0.96	0.94	0.86	0.91	0.91	0.94
$p_0^{(h)}$	[-]	0.88	0.88	0.87	0.89	0.88	0.86	0.83	0.83	0.83	0.87
$p_0^{(d)}$	[-]	0.55	0.52	0.49	0.56	0.51	0.48	0.49	0.48	0.46	0.50
$r_1^{(5min)}$	[-]	0.80	0.61	0.63	0.74	0.69	0.63	0.87	0.73	0.66	0.70
$r_1^{(h)}$	[-]	0.44	0.41	0.37	0.41	0.39	0.40	0.51	0.48	0.47	0.38
$r_1^{(d)}$	[-]	0.16	0.21	0.17	0.20	0.18	0.13	0.25	0.28	0.20	0.22
$H_{max}^{(5min)}$	[mm]	11.04	21.1	11.44	10.7	11.7	13.31	11.53	17.99	11.49	14.3
$H_{mean}^{(5min)}$	[mm]	0.09	0.20	0.13	0.08	0.18	0.11	0.09	0.14	0.12	0.12
$H_{std}^{(5min)}$	[mm]	0.21	0.30	0.27	0.18	0.31	0.25	0.16	0.23	0.24	0.26
$H_{skew}^{(5min)}$	[-]	13.98	16.87	12.35	17.09	14.20	18.04	12.54	16.52	13.68	15.65
$H_{max}^{(h)}$	[mm]	34.19	32.7	27.46	39.89	48.1	34.08	37.14	46.54	24.91	37.3
$H_{mean}^{(h)}$	[mm]	0.85	0.89	0.75	0.67	0.72	0.60	0.83	0.83	0.78	0.68
$H_{std}^{(h)}$	[mm]	1.46	1.42	1.35	1.19	1.41	1.18	1.31	1.35	1.34	1.29
$H_{skew}^{(h)}$	[-]	6.97	5.96	7.11	9.53	11.02	8.86	7.33	8.22	5.68	7.70
$H_{max}^{(d)}$	[mm]	70.63	49.6	60.6	60.89	55.1	46.27	84.2	69.24	58.57	62.62
$H_{mean}^{(d)}$	[mm]	5.50	5.36	4.49	4.11	4.08	3.89	6.72	6.64	6.03	4.13
$H_{std}^{(d)}$	[mm]	6.96	6.69	6.08	5.39	5.81	5.48	8.21	8.45	7.82	6.00
$H_{skew}^{(d)}$	[-]	2.92	2.32	2.90	3.16	3.31	2.98	2.83	2.55	2.29	3.26

Table 3.2.: Overview of some basic global statistics of the pluviometers in Freiburg (F.), Weiden (W.), Kempten (K.) and Augsburg (A.) for three time periods (I – III). *Remark:* H_{daily} is the data from a daily rainfall gauge and not of from the high resolution pluviometer.

the sum of the respective daily rain gauge ΣH_{daily} . Therefore, most of the missing data must be related to missing zero precipitation values. The reason is that data of period I was derived from digitized paper chart records for which a digitalization was only performed if precipitation occurred. Unfortunately, I have no additional information if the nan-values are actual missing data or if they are not digitized time periods. Therefore, I treat all nan-values as zero values in all three periods. Although this procedure underestimates the true local precipitation for some periods, it matches the approach that is usually applied in practice when rainfall is used in sewer simulations. Furthermore, the absolute precipitation values are not too important for my study for two reasons. First, I mainly focus on the temporal structure rather than on the rainfall distribution. Second, I do not apply the data for real case sewer simulations to investigate critical system conditions. Instead, I use them as references for relative comparisons with simulated time series. Consequently, in the rest of the study, that is for the statistical analysis and sewer simulations, I set all missing values to zero and use these series as references. Nevertheless, one should keep in mind that for a full simulation of synthetic time series one should take care of missing values, for example, by using the more robust daily stations for statistics where applicable, or by applying infilling methods (e.g. *Bárdossy and Pegram, 2014*).

3.3. Influence of different measurement devices on precipitation statistics

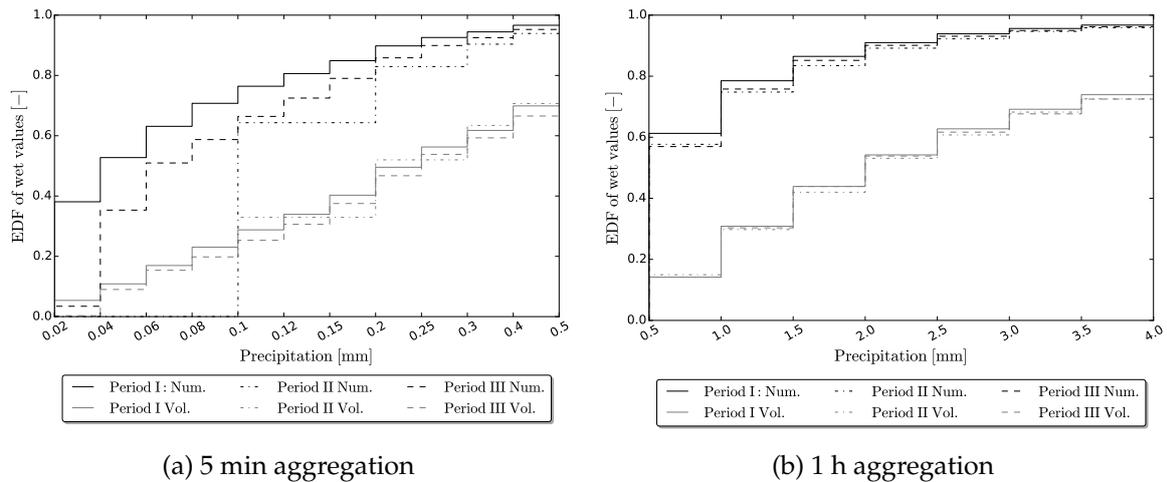
Systematic differences in the statistics can be observed in Tab. 2.1 when different time periods are regarded. Some of the deviations can be related to sampling errors of short time periods. For example, the low mean annual precipitation sum of Freiburg for time period III can be referred to rather dry years between 2007 and 2012. However, not all deviations need to be caused by sampling errors. Some differences could be related to natural variability in the climate. However, due to the small sample sizes this hypothesis can neither be rejected nor proofed.

Other deviations are likely induced by different measurement devices. As no secondary information about the devices are available for my study, I can also not prove this hypothesis. Nevertheless, I present plenty of empirical evidences in the following that indicates a strong influence of different devices on certain rainfall statistics. They can already be noticed by a visual analysis of the time series (Fig. 3.1).

Period I reflects data of subsequently digitized paper chart records. The digitalization of paper records is done by locating points that detect remarkable inflections of the registered curve (*ATV-A136, 1985*). The points are then connected with straight lines such that the traverse fits the precipitation record. The rainfall depth is defined as difference between two points in 0.01 mm. As in the digitalization the increase of precipitation depth is performed linearly, the resulting intensity between two points is constant. That is, the representation of the temporal structure of the intensity depends on the amount of remarkable points and, thus, on the quality of digitalization. In particular for small intensities, the amount of remarkable points can be very small. The result is a very smooth time series and

Abbreviation	Time period	Resolution [mm]	Characteristics
Period I	196x – 198x	0.01	smooth intensity, many small values, few zeros
Period II	199x – early 200x	0.1	sharp spikes, many zeros
Period III	late 20xx	0.01 – 0.03	mix between period I and II, most realistic

Table 3.3.: Overview of the characteristics of different periods used in this study.



(a) 5 min aggregation

(b) 1 h aggregation

Figure 3.2.: EDF of the number and the respective volume of the wet time series ($H > 0$) at 5 min and 1 h aggregation of Freiburg for three time periods (I – III). Remark: Because the 5 min data is highly skewed, the EDF of (a) is classified not equidistantly.

is the smoother, the smaller the precipitation volume and the larger the respective time delta is.

Period II corresponds to automated pluviometers of the first generation of the DWD with a coarse resolution of 0.1 mm. The characteristics are typical for tipping buckets, where the precipitation accumulates to 0.1 mm before an impulse is detected due to tipping. When the precipitation intensity is less than 0.1 mm/5 min, the precipitation depth is recorded as 0 mm until the accumulated rainfall exceeds 0.1 mm. The resulting time series shows single spikes of 0.1 mm interrupted by 0 mm values.

Since period III the pluviometers from the *Automatische Messdatenerfassungsanlage* [automated gauging station] (AMDA) have been available. These pluviometers have an increased resolution up to 0.01 mm and can be, for example, weighing pluviometers. As the resolution is higher, even small precipitation intensities can be resolved well. The resulting time series has a more detailed intensity structure compared to period II, and is less smooth than period I due to the obsolete linear interpolation between two signals.

Table 3.3 gives a comprehensive overview of the characteristics of the sub-periods. The intensity structure ranges from highly smoothed (period I) to very scattered (period II). The

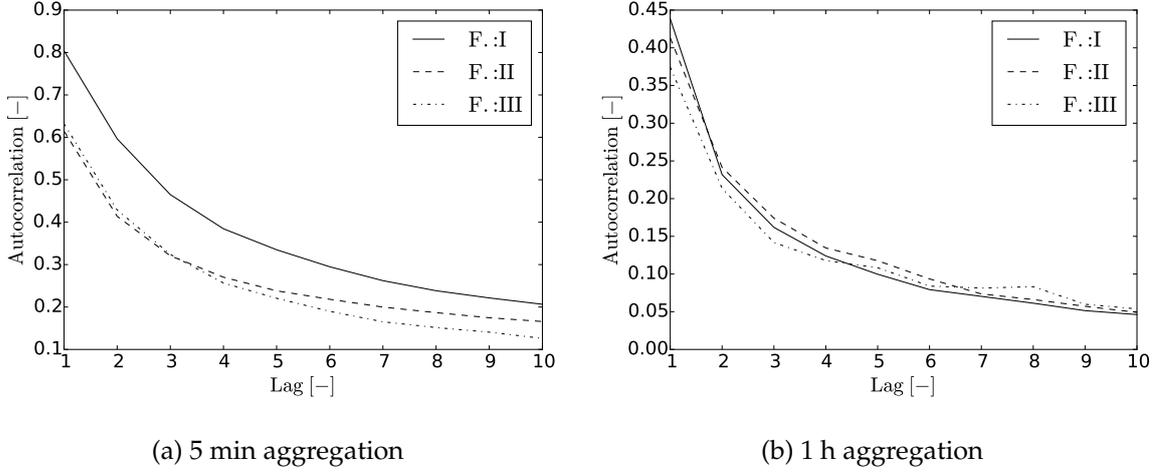


Figure 3.3.: The autocorrelation function (ACF) of different aggregations of Freiburg for three time periods (I – III).

different characteristics of period I and II compared to the more realistic period III are mostly caused by small 5 min precipitation intensities. The larger the intensities, the more the differences reduce as illustrated by the EDF in Fig. 3.2a. This can be easily explained if one reconsiders the measurement procedures. In period I, the time between two remarkable points decreases with increasing intensity, and thus the number of values with constant intensity reduces. The result is a less smoothed time series that is more similar compared to period III. For period II, larger intensities exceed the resolution of the tipping bucket device at each time step, without having a pre-filled bucket that tends to tip. Additionally, for increasing intensities the relative contribution of each single tip to the total intensity of one time step reduces. Thus, the coarser resolution of period II compared to period III is less relevant.

Although the differences mainly effect small precipitation events, the impact on certain statistics can be very high. As already mentioned, large deviations occur in the cumulative frequencies of the wet 5 min values (Fig. 3.2a). The figure indicates a larger number of very small 5 min values in period I due to the smoothing effect of small events. Another example is $p_0^{(5min)}$, which is the smallest for period II due to the discrete values of the tipping bucket and highest for period I (Tab. 3.2). Respectively, the number of wet 5 min values $n_{wet}^{(5min)}$ is the highest in period one (Tab. 3.2). Both characteristics are also caused by the smoothing effect of small events. For the same reasons, the autocorrelation $r_1^{(5min)}$ is increased in period I (Fig. 3.3a). The scaling parameters are also highly effected, but I will discuss them in more detail later in Chap. 3.4.

As soon as larger aggregations are regarded, the systematic deviations in the EDFs, in p_0 and in $r_1^{(j)}$ disappear, or at least clearly diminish (see Fig. 3.2b, Tab. 3.2 and Fig. 3.3b respectively). Other statistics, however, still exhibit a considerable influence. An example is the sum of the frequencies of wet spells in hourly resolution $ws_{\Sigma}^{(h)}$, where the spells are separated by at least one dry hour. Table 3.2 clearly indicates a systematic underestimation

of $ws_{\Sigma}^{(h)}$ for time period I, although the statistic is calculated at the hourly resolution. The underestimation of the wet spells is mostly caused by the reduced number of short spells ($ws_{1h}^{(h)}, ws_{2h}^{(h)}$). The underestimation is remarkable, as the number of wet hours $n_{wet}^{(h)}$ is only slightly reduced in period I compared to the other time periods.

One can conclude that the influence of the measurement devices on statistics should be considered depending on the purpose and on the regarded time scale. It is of minor relevance as long as applications are regarded, where the dominant factors are the integral precipitation volume, the intensity of large precipitation events or highly aggregated time series. However, as soon as time series of high temporal resolution are statistically analyzed, for example, for the generation of synthetic precipitation, the different measurement devices play an important role. On the other hand, one should keep in mind the sampling errors of different and short time periods that can additionally influence the statistics.

The conclusion of this section is that all statistics calculated in a high temporal resolution should be analyzed individually towards systematic inhomogeneities between time periods of different measurement devices. If deviations are observed, further investigations must be performed with respect to their influences on subsequent applications. If large impacts are detected, I recommend to use only high resolution data of period III for vulnerable statistics, because they reflect the natural intensity structure most realistically. For statistics not affected, all data should be applied to reduce sampling errors of short periods.

3.4. Issues using scaling parameters

The scaling parameters are an important set of parameters that is used later for the generation of synthetic precipitation. They are investigated in more detail now as several issues occur related to the measured data. The scaling parameters are used to describe the temporal structure of precipitation across a large range of aggregations and can be obtained for different moments. The higher the order of the calculated moments, the more they are affected by large values, and the more vulnerable they are to outliers. Consequently, only the lower order moments should be considered for any application. For example, *Lombardo et al.* (2014) recommends to use only the first two moments. In the following, the first three moments are investigated as they are used for the generation of synthetic precipitation in NiedSim 2.x. NieSim 2.x is the operational version of the precipitation simulator, on which my study is build on. For more detailed information about the algorithm of NiedSim 2.x, I refer on Chap. 4.4.

In order to obtain the scaling parameters of the first three moments, the moments $m_k^{(j)}$ with order $k = \{1, 2, 3\}$ are calculated for the aggregations $j = \{5, 10, 20, 30, 60, 120, 240, 360, 720 \text{ and } 1440 \text{ min}\}$ according to Eq. 2.2. The scaling parameters b_k are then obtained using Eq. 2.3. They are calculated for summer and winter individually as the scaling characteristics change in different seasons. In NiedSim 2.x summer is defined from May to October and winter from November to April.

In order to interpret the scaling parameters, the power function has to be recapitulated. A scaling parameter $b_k = 1$ dedicates a linear relationship between the moment k and the

aggregation level j . The scaling parameter of the second moment b_2 of precipitation shows roughly a linear relationship.

If $b_k > 1$, the relationship of the moment k and aggregation level j is convex. This is the case for the third moment b_3 of precipitation. The more uniform the values of the high resolution are in comparison to the low resolution, the smaller m_3 of high resolutions compared to low ones is, and thus the larger the scaling parameter b_3 is.

For scaling parameter $b_k < 1$ the dependence between the moment k and the aggregation level j is concave. The first moment, the mean, must follow this characteristic as one can easily imagine. If b_1 was not < 1 , but for example $b_1 = 1$, a small 5 min value of 0.1 mm (the equivalent of a single tip of tipping buckets) would result in a large daily precipitation value of 28.8 mm (a rainfall depth, that occurs approximately once per year).

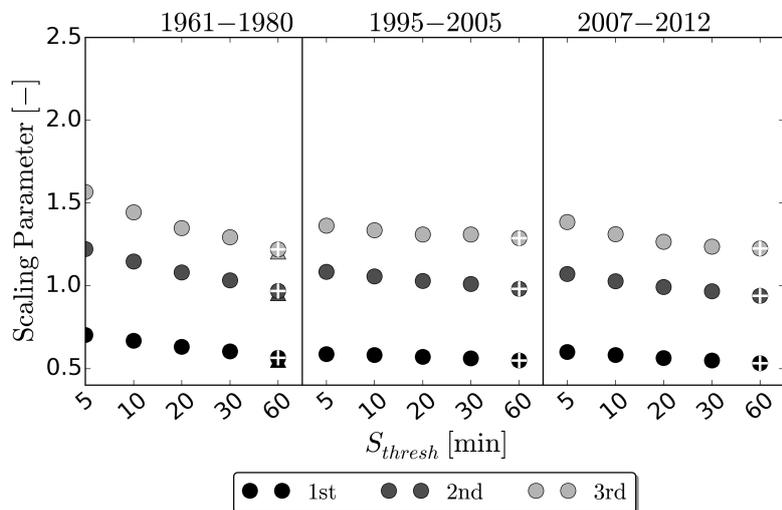
3.4.1. Breaks in scaling property

In NiedSim 2.x, all aggregations from 5 min to 24 h are used to calculate one set of scaling parameters. However, as for example *Hoang et al.* (2012) shows, scaling breaks occur for sub-hourly precipitation data. In order to investigate the breaks, I calculate the parameters with different subsets of aggregation levels for which the moments are computed. I define two cases of subsets. Case (a) includes only aggregations that are larger or equal a threshold aggregation S_{thresh} . Consequently, the subsets of case (a) result in aggregation levels from S_{thresh} to daily aggregations S_{24h} . This case pronounces the scaling behavior where the focus is on highly aggregated data. Conversely, in case (b) the subsets are based on aggregations less or equal S_{thresh} and start from the 5 min aggregation. Therefore, the intra-hourly aggregations dominate the scaling parameters in case (b).

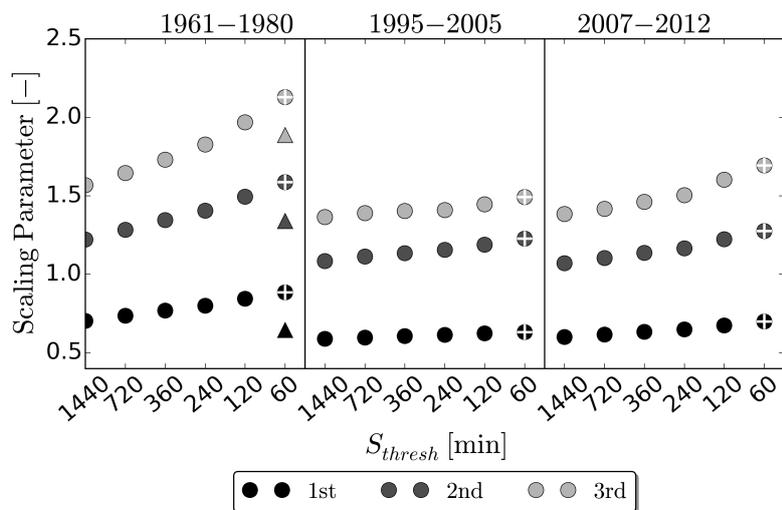
Fig. 3.4 shows the results of both cases for three different time periods in the summer season. The parameters to the very left of each time period of (a) and (b) are identical ($S_{thresh} = 5$ in (a) and $S_{thresh} = 1440$ in (b)), because they are based on the entire range of aggregations. The number of considered aggregations decrease if one moves towards the right side of each time period. In both cases, the subsets are limited to $S_{thresh} = 60$ min. Otherwise, too few aggregations would be used for the regression and would result in very uncertain scaling parameters. If the scaling behavior of the moments was the same from 5 min to daily aggregation, the scaling parameters would be constant across the scales. However, a clear change can be observed if not the entire range of aggregations is used. Independent of the time period, the parameters tend to decrease the more large aggregations are considered (a), and tend to increase if the focus is on high resolutions (b).

3.4.2. Sensitivity to measurement devices

The scaling parameters also show systematic deviations if they are calculated with different time periods (Fig. 3.4). The more high resolution data is used, the larger the differences are (Fig. 3.4b, in particular for 1961 – 1980). The deviations might be caused by sampling errors of different sample sizes. However, an analysis of subsamples of equal size exhibited



(a) Aggregation levels $\geq S_{thresh}$ are used: $S_j = [S_{thresh}, \dots, 1440]$



(b) Aggregation levels $\leq S_{thresh}$ are used: $S_j = [5, \dots, S_{thresh}]$

Figure 3.4.: Scaling parameters of the moments $k = \{1, 2, 3\}$ calculated with different subsets of aggregations S_j . Each subset is defined by the threshold aggregation level S_{thresh} . The figure shows the scaling parameters in summer of three time periods of Freiburg. The marked dots represent $S_{thresh} = 60$ min which is the aggregation level that separates the 1 h optimization and 5 min disaggregation of NiedSim. The triangle reflects the data set transformed with a synthetic tipping bucket emulator.

the same deviations as the full samples. There are very convincing indicators that different measurement devices are again responsible for the systematic deviations as I will show in the following.

In Chap. 3, I demonstrated that the influence of the devices is highly reduced at the hourly aggregation. If only aggregations larger than 1 h are used for the scaling parameters (crossed circle in Fig. 3.4a), almost no deviation between the time periods can be observed. On the other hand, large deviations can still be observed if only aggregations smaller than one hour are applied (crossed circle in Fig. 3.4b). Additionally, the smoothed time series of period I (1961 – 1980) shows systematically larger scaling parameters of all moments compared to the other two time periods.

To support the assumption, I apply a simple tipping bucket emulation algorithm to the time series of 1960 – 1981. That is, I use the smooth 5 min time series in 0.01 mm resolution as input of a synthetic tipping bucket emulator with 0.1 mm resolution. For each time step, the new intensity is calculated as a multiple of 0.1 mm. The reminders are accumulated in a synthetic tipping bucket until 0.1 mm is reached. I plot the scaling parameters of the resulting time series as triangles in Fig 3.4. The parameters of all moments are systematically reduced and are more similar compared to the other periods. This illustrates the high influence of measurement devices on scaling parameters which was also reported by *Hoang et al.* (2012).

All problems related to the scaling parameters are also prominent for the winter season as well as for the station Kempten and Weiden (results not shown) and support their general validity.

3.5. Issues with extreme values

Finally, I would like to address briefly the large and extreme values of the observed data used in my study. For this purpose, I calculate the partial series as explained in Chap. 2.2.3. I use the same number of events as defined in KOSTRA, that is, the Euler's number $e \approx 2.718$ times the number of years. Figure 3.5 illustrates the IDF-curves for two different durations (5 min and 30 min).

The erratic bends in the lines for large return periods indicate the high sampling uncertainty of extreme events. The length of measured data should be at least three times the desired return period. This is a rule of thumb based on some theoretical considerations using the binomial distribution. The probability of occurrence of an event in a time period of 20 years with a return period of 20 years is only 64 %. If an event with a shorter return period of, for example, 6.6 years is regarded, the probability increases to 96 % in a 20 year period.

Therefore, if one would like to predict extremes of long return periods, one either needs very long observed time series that are usually not available, or one has to fit theoretical extreme value distributions. One can easily fit such distributions, but due to the lack of

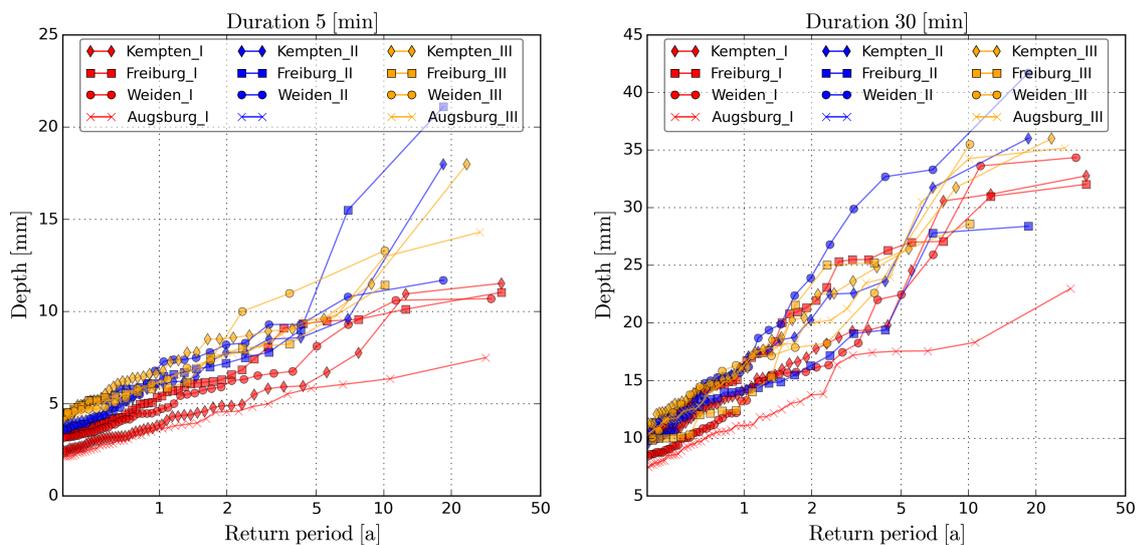


Figure 3.5.: Partial series of four locations and three time periods for the duration 5 and 30 minutes.

missing observed data, the validation is very difficult and thus the results are rather uncertain. Therefore, the most limiting factor with respect to precipitation extremes are the short observed time series of high resolution rainfall data.

For the following considerations, I focus on the short return periods as for return periods < 1 year the sampling uncertainty is rather small indicated by little fluctuations of the graphs in Fig. 3.5. A potential issue may occur if temporal resolutions close to the resolution of the measurement device are regarded. If one considers the partial series at a duration of 5 min (Fig. 3.5 left panel), one can observe significantly smaller precipitation depths of period I compared to time period II and III at all locations. One might argue that short intense rainfall has been increased over the years due to changes in the long term precipitation characteristics. However, the effect diminishes for higher aggregations as illustrated in Fig. 3.5 (right panel) showing a duration of 30 minutes. Bearing in mind the influence of different measurement devices on certain statistics, it seems to be more likely that they also explain the systematic deviations in the partial series at short durations. However, as I had no overlapping time periods, I could not verify this assumption.

4. Synthetic precipitation

4.1. Motivation

Precipitation is one of the major input variables for applications in urban hydrology. Depending on the specific application, the precipitation must meet different requirements. Three major applications can be defined in urban drainage: hydraulic design of sewer networks, planning of combined sewer systems and sewer system management. However, to be able to understand the requirements with respect to rainfall, I give a short overview here: For hydraulic design, extreme values are most relevant, whereas for combined sewer systems as well as for sewer system management, the mean yearly conditions are important including the temporal structure of medium precipitation events. In the following, basic properties of precipitation time series are explained that define the minimum requirements for sewer modeling.

Length The length of measured time series is a crucial aspect that must be considered, in particular, for the calculation of extreme values. As extreme values by definition occur only rarely, it is likely that short periods will miss large events. A rule of thumb as defined in Chap. 3.5 is that the measured period should be at least three times as long as the considered return period. In the literature, the demand is less strict depending on the application. *DWA-A531* (2012) states that the period should be at least 30 years such that local extreme value statistics can be properly assessed. Whereas according to (*DWA-A118*, 2006), the measured periods should be at least ten years to reflect the return periods from 1 to 10 years for design purposes.

Short measured time periods do not only affect extremes, but also other statistics due to sampling errors. The higher the year to year variability, the larger the sampling uncertainty of short measured data is. For example, the annual precipitations sums can highly differ from one year to another and, thus, are vulnerable to short periods. As explained in Chap. 3.2, the mean annual precipitation of the short period III of Freiburg is reduced compared to the long term mean due to the rather dry years. Consequently, even for applications, where the extremes are less important like for combined sewers, long time series are required (according to *ATV-A128* (1992) at least ten years). In conclusion, long time series are necessary to avoid sampling errors for most sewer applications.

Continuity Sewer simulations can be performed either event-based or continuously. Event based simulations are discontinues because relevant single precipitation events with typical

characteristics are selected, individually simulated and analyzed. The single events can either be selected from measured data or generated as synthetic events. An example for the latter case are design storms used for the hydraulic design of sewer networks (*DWA-A118*, 2006). The intensity and durations of such synthetic events are based on tabulated IDF values like the KOSTRA data, and their internal intensity can be uniform (block rainfall) or follow an artificial distribution, for example Euler Type II (*DWA-A118*, 2006).

Other applications, for example tasks related to combined sewer systems, require a realistic representation of wet and dry periods. Several studies showed that event-based simulations cannot account for such alternating characteristics and lead to large errors compared to continuous time series (*Xanthopoulos*, 1990; *Drechsel*, 1991). In such cases, continuous simulations have to be performed for which measured time series must be available.

Another aspect is the treatment of data gaps that usually occur in observed data sets. Even for well maintained pluviometers, missing values occur that are caused by different problems in the measuring technique. For many applications the missing values can be supposed to be zero if the number of missing values are small. This treatment, however, will most likely lead to an underestimation of the precipitation depth unless all missing values occurred in dry periods. If the assumption is acceptable, and if the introduced error is negligible, must be checked by the user in advance.

Temporal resolution Another subject is the temporal resolution of the time series. The necessary resolution depends on the size of the regarded catchment and thus on the drainage time. Most applications concentrate on catchments of a few to a few hundred km². In catchment hydrology, the drainage times are in the order of hours or days, and thus hourly or even daily time series are sufficient. In the context of urban hydrology, the response time of the system is much faster due to shorter drainage times mainly related to a high percentage of impervious areas. Consequently, a high temporal resolution in the range of several minutes can be necessary as shown for example in *Schilling* (1991); *Berne et al.* (2004); *Ochoa-Rodriguez et al.* (2015). Precipitation data from daily measurements are available in Germany for more than 100 years. The length of high temporal resolution series, however, is significantly shorter compared to daily data and 5 min time series with more than 50 years of data are very rare.

Spatial availability Unfortunately, only few locations exist for which the time series fulfill all requirements. This data could be sufficient if the characteristics of precipitation were spatially uniformly distributed. However, precipitation can show a high spatial variability related to specific local topographical and climatological conditions. Thus, even if the annual rainfall depth of two stations are similar, several other statistics like extremes or dry duration periods can be different, and their consequences on subsequent application are uncertain (*Schilling*, 1991). Consequently, time series applicable for urban hydrology cannot be simply transferred to any other location. Alternatives are needed for ungauged areas or for locations where the requirements are not fulfilled.

4.2. Alternatives to measured time series

The concept that is usually applied to obtain data for an unknown location is interpolation. Based on a few observed data sets in space, the data between the observations is estimated. In the case of precipitation time series, however, interpolation is no solution as it overestimates the autocorrelation, underestimates the annual maximum as well as the variance of precipitation values and leads to reduced extreme values (*Bárdossy et al., 2000*). If a spatially interpolated time series was applied in urban hydrology, the result would be an underestimation of the failure probability in design tasks or in wrong quality assessments regarding wastewater treatment problems.

Another idea is to take a measured time series of good quality of one location and transform it in such a way that it is representative for the local conditions of another station. A transformation could be, for example, a rescaling of the time series with the mean annual precipitation sum. This is applicable for specific tasks and some regions as shown by *Drechsel (1991)* for combined sewage systems in Hesse (Germany). Such an approach assumes the same temporal structure of precipitation values for the gauged and ungauged location. The assumption can be validated for certain statistics or specific applications, but a general verification is very difficult.

A different solution for unobserved locations can be the generation of synthetic precipitation based on interpolated statistical parameters representing the local rainfall characteristics. The advantage is that such time series can be generated for any location, are continuous and without missing values, can be of arbitrary length and in a high temporal resolution. Compared to the interpolation, the statistics, in particular, the variance and the extremes, can be preserved. Another advantage is the possibility to generate different samples, at least if the time series are generated with a stochastic simulation. In such cases, uncertainty estimations can be evaluated and risk assessments performed. Furthermore, if synthetic time series are based on a well elaborated regionalization approach, the synthetic time series of different locations are homogeneous in the way they are generated, and so are better comparable with each other. Except for difficulties when setting up the regionalization including the estimation of parameters, no additional sampling errors, measurement errors or quality problems exist. Additionally, synthetic time series at different locations are not influenced by varying measurements techniques or different measured time periods. Finally, a good synthetic time series can fulfill all requirements on precipitation data, and thus would not be restricted to special applications. Nowadays, due to the increased computational power, this is a major advantage as different tasks can be simulated simultaneously, for example, sewer surcharging and combined sewer overflow simulations.

Several approaches (parametric and non-parametric, deterministic and stochastic) are available to generate precipitation as explained in the following section.

4.3. Overview of different precipitation models

Many different approaches for the generation and disaggregation of continuous precipitation time series have been developed over the last decades. In the context of precipitation modeling, the term *generation* of precipitation describes the construction of a sequence of events or a series in a discrete resolution. Whereas the term *disaggregation* refers to a method that refines an existent time series at a certain temporal resolution to a higher resolution. Most synthetic precipitation time series are generated at a daily or hourly aggregation and then, in a second step, are disaggregated to hourly or sub-hourly scales.

Only a few deterministic precipitation models exist, for example, *Ormsbee* (1989) who developed a model where each intra-hourly distribution is defined depending on the shape of the previous, the actual and the following hour. The majority of synthetic precipitation models are based on stochastic processes. The reason for most models being stochastic is that atmospheric processes are highly complex and cannot be described completely deterministically. Instead, all atmospheric processes are combined in and reflected by a stochastic process where each observed series is a realization of the underlying population. Unlike deterministic models, stochastic models have not a specific output for a set of input parameters, but a distribution of outputs describing their probability. Consequently, a stochastic description of precipitation includes an implicit evaluation of the uncertainty.

Some of the earlier stochastic approaches for rainfall modeling are based on chain dependent processes. Such a model is separated in an occurrence process described by a Markov Chain (MC), and an intensity process modeled for example by an autoregressive process. In a MC, the state (rainfall or no rainfall) of a time step depends on the state of one (first-order MC) or more previous time steps (higher-order MC). The change of the state is described by transition probabilities. These models can easily account for non-stationarity like seasonal or diurnal cycles (*Katz and Parlange, 1995*). However, as they cannot account for higher order statistics or scaling behavior (*Paschalis et al., 2014*), these models are mostly applied to daily precipitation models, in particular, for weather generators (*Wilks and Wilby, 1999*). For the representation of rainfall at high temporal resolutions other models are usually applied. The three most common types are Alternating Renewal models, Point Process models and Multiplicative Random Cascade models.

Alternating Renewal models are a further development of the MC approach. Similar to the MC, the precipitation occurrence and precipitation intensity are modeled separately. In Alternating Renewal models, however, the time series is composed of complete precipitation events instead of values at single time steps. Each event consists of a wet and a dry period and is described by an external (occurrence process) and internal structure (intensity process). The external structure is fully described by the duration of wet and dry spells. The internal structure is defined by the volume or intensity within each wet spell. All three parameters are characterized by distribution functions that cannot be modeled independently of each other. Several approaches exist to describe the dependence structure, for example: simple models, where the wet spell intensity and duration are described with a joint probability distribution (*Robinson and Sivapalan, 1997*); Generalized Linear Models using non-linear (logistic) regression (*Chandler and Wheeler, 2002*); complex models including the full

dependence structure between all three event characteristics using copulas (*Salvadori and Michele, 2006*). The internal structure is often modeled independently of the external structure and defined by a distribution function of the precipitation intensity. The distribution can be described in several different ways, for example by a double exponential function with a defined peak (*Haberlandt, 1998*), by Fractional Noise (*Bernardara et al., 2007*) or by mass curves (*Robinson and Sivapalan, 1997*). An advantage is that Alternating Renewal models rely only on a few parameters that all can be measured directly. However, the model is based on the definition of wet and dry spells, in particular, on the definition of the separation time. This crucial step for the model performance is not unique, and thus is a major disadvantage of the model. Several different definitions exist depending on statistical considerations as well as on subsequent applications (see Chap. 2.2.2). Another problem of these models is the tendency to underestimate the autocorrelation (*Paschalis et al., 2014*).

The second common group of models are Point-Process models where at certain points in time (rectangular) pulses emerge that represent single rain cells. In the simplest version, the time between two consecutive pulses is Poisson distributed. The single cells are parametrized with depth and duration. The total precipitation depth at a specific point in time can then be calculated by the summation of all individual cells active at that time. In more sophisticated approaches, at first storm cells of variable durations are defined, where the occurrences follow a Poisson process. In each storm cell, several individual rain cells can develop that are described in two different manners. In the Bartlett-Lewis model, the waiting time between two individual rain cells are independent and identically distributed, whereas in the Neyman-Scott process, the positions of the rain cells with respect to the corresponding storm cell are independent and identically distributed (IID) (*Rodriguez-Iturbe et al., 1987*). Like in Point Based models, each single cell is defined by depth and duration, and the precipitation depth at each time step is the sum of the individual cells. Due to the model structure, the Point-Process models usually represent the autocorrelation better compared to other models (*Paschalis et al., 2014*). Another advantage is the independence of the applicable time scales from the input parameters (*Koutsoyiannis and Onof, 2001*). Furthermore, the model can be easily transferred to spatial scales by using discs instead of points (*Cowpervait, 1995*). Although Point Process models can be related to physical processes of storms (*Onof et al., 2000*), the actual single pulses cannot be observed in nature. Therefore, the calibration is complex and can only be done indirectly using statistical characteristics of the time series for example the autocorrelation, dry periods or moments of the precipitation distribution (*Rodriguez-Iturbe et al., 1987*). Additionally, these models often show a good performance for the calibration scales, but perform weak for other aggregations, in particular, for high temporal resolutions at sub-hourly scales (*Paschalis et al., 2014*). Other draw backs are the common underestimation of extremes and p_0 values (*Pui et al., 2012; Verhoest et al., 1997*).

Multiplicative Random Cascade models, the third group, are based on the theory of self-similarity and scale invariance. These models are often applied for downscaling of precipitation. The high resolution time series is obtained via a branching system of an initial series at a coarser resolution. At each aggregation the precipitation amount at the coarse scale is split by a so called cascade generator into separate parts (branches). That is, the precipitation structure at a higher resolution develops from the existing structure at the coarse aggregation and is defined and calibrated by the cascade generator. An advantage of cas-

cade models is that a high temporal resolution model can be set up with coarser calibration data (*Koutsoyiannis and Onof, 2001*) or with data sets of different temporal resolutions. Additionally, the models are parsimonious, that is, only very few parameters are necessary to set up the model (*Molnar and Burlando, 2005*). According to them, the random cascade model can also be easily extended to a spatial dimension. On the other hand, these models show problems in the representation of the autocorrelation as shown for example by *Müller and Haberlandt (2016)*; *Pui et al. (2012)*. Furthermore, the model is very sensitive in the calibration, in particular, if high moments are considered and, thus, is highly affected by sampling errors (*Onof et al., 2005*; *Gaume et al., 2007*). Finally, the scaling process of natural rainfall may not be the same for all temporal aggregations from 5 min up to daily precipitation as shown in Chap. 3.4.1. Several papers underline occurring scaling breaks at higher temporal resolutions (e.g. *Fraedrich and Larnder, 1993*; *Molini et al., 2009*).

Further, but less common models are available for the disaggregation. *Burian et al. (2000)* used artificial neural networks, in which the information at the coarse aggregation is propagated to the fine scale via interconnected hidden layers. Although such models can be fit to relate any kind of input and output data, the actual processes are hidden in a black box and can neither be analyzed individually nor observed empirically. Another concept is the method of fragments, a resampling approach, where daily values are disaggregated by sub-hourly fragments of nearby high resolution observations of similar daily properties (*Westra et al., 2012*). According to them, a major drawback is the reduced interconnectivity of the temporal patterns at the daily aggregation.

Each model described above shows some advantages, but also suffers from weaknesses. Therefore, recent approaches combine several model types to overcome the disadvantages of the individual models. For example, *Paschalis et al. (2014)* simulated daily time series with an external model and performed the disaggregation with a separate internal model. They analyze different combinations of external models (Alternating Renewal, Markov Chain and Point Process) and internal models (Multiplicative Random Cascade). They show that the combined models performed better compared to the individual models, but even the combined models were not able to reproduce the autocorrelation correctly.

4.4. The NiedSim approach

The approach I investigate in this study is based on Markov Chain Monte Carlo as well as Simulated Annealing techniques and was developed by *Bárdossy (1998)*. It is a data-driven, non-parametric stochastic precipitation generator. In contrast to other approaches, no physical process is modeled or described stochastically. The basic concept is to optimize the pure statistical properties of the time series such that the synthetic and observed series cannot not be distinguished in their statistics. It assumes that certain statistics can fully describe the natural characteristics of rainfall. Since the first paper, subsequent research using this concept has been dedicated including approaches towards simultaneous modeling of spatio-temporal time series (*Brommundt, 2008*; *Beck, 2012*).

The operational version for the generation of single point time series is called *NiederschlagsSimulator* (NiedSim) [precipitation simulator] and has been used by four Federal States of Germany for several years: in Baden-Württemberg since 2000 (*Bárdossy et al.*, 2000), in Hesse and in Rhineland-Palatinate since 2002, in Bavaria since 2008. In 2014/15, the input data was updated with more recent measurements and the controlling software was changed. The updated version is called NiedSim2.x and is the basis of my study. Issues in the performance of NiedSim occurred after its update with new measured data which was one of the motivations to investigate NiedSim again. The other aspect was to improve the algorithm of NiedSim with a larger focus on its influence on subsequent sewer simulations. The innovation of my work is the generation of precipitation along with its indirect validation using a sewer model. Such an iterative approach enables to detect issues in synthetic series more easily. Furthermore, potential errors can be evaluated with respect to their relevance for sewers and finally, they can be corrected to improve the quality of synthetic time series.

NiedSim can be subdivided into three major parts which work independently from each other and, thus, can be analyzed individually. The first part is the database that contains all observed precipitation statistics. The second part performs the generation of a continuous, hourly precipitation time series. In the final part, the hourly time series is disaggregated to a time series in 5 min resolution. NiedSim is mainly programmed with the open source language Python 2.7 (*Python*, 2015) including the generation of the database and the controlling of the individual routines. However, the Simulated Annealing algorithms of the optimization routines are coded in Fortran due to computational speed benefits. In the following, the different parts of NiedSim2.x are explained in more detail.

4.4.1. Database

The basis for NiedSim are observed point precipitation data at different temporal resolutions. Daily, 1 h and 5 min time series are used to calculate various statistical characteristics of precipitation. Table 4.1 gives an overview of the statistics used in NiedSim2.x. I explained most statistics in Chap. 2.2 except for the cross-correlation, the weighted autocorrelation and the circulation patterns (CPs) because they are rather specific for NiedSim.

The cross-correlation represents the correlation of the time series of the target location with a predefined reference time series and should preserve some spatio-temporal continuity if more than one station is simulated.

The weighted autocorrelation is introduced to account for seasonal variations in the autocorrelation. It is calculated in the same way as the autocorrelation, except that the precipitation values are seasonally weighted before applying Eq. 2.1. A sine function is used to reflect a continuous change of weights between 0 and 2 during the year. The larger the weight, the more influence the respective season on the weighted autocorrelation has.

CP statistics describe classified precipitation characteristics related to synoptic pressure systems. They were introduced by *Bárdossy and Plate* (1992) to generate daily precipitation time series, such that large scale weather characteristics can be considered.

For more detailed information about the applied parameters in NiedSim2.x, it is referred to *Bárdossy et al. (2015)*.

Parameter	Calculation period	Aggregation [min]
1 h generation		
ϕ^* : hourly distribution	whole period	60
ϕ : monthly sum	monthly	-
ϕ : daily maximum	annual	1440
ϕ : daily exceedance probability of 0, 1, 5 mm/day	annual	1440
ϕ : CP statistics (mean, std, probability)	whole period	1440
ψ : autocorrelation (lag 1)	annual	60, 120, 180, 360, 720, 1440
ψ : weighted autocorrel. (lag 1)	annual	60, 120, 180, 360, 720, 1440
ψ : scaling exponent (1 st – 3 rd moment)	winter / summer of whole period	5, 10, 20, 30, 60, 120, 240, 360, 720, 1440
ψ : cross-correlation	whole period	60, 1440
5 min disaggregation		
ϕ : 5 min maximum	annual	5
ϕ : 1 h maximum	annual	60
ϕ exceedance probabilities (0.45, 0.95, 1.95 mm)	whole period	5
ψ : autocorrelation (lag 1,...10)	whole period	5
ψ : weighted autocorrel. (lag 1)	whole period	5
ψ : scaling exponent (1 st and 2 nd moment)	winter / summer of whole period	5, 10, 20, 30, 60, 120, 240, 360, 720, 1440

Table 4.1.: Parameters used in NiedSim2.x for the 1 h generation and 5 min disaggregation. ϕ^* is only used for the setup of the initial time series and not for the optimization.

It should be highlighted that according to the definitions of Chap. 2.2 only global and extreme value statistics are used in NiedSim, and no event-based statistics. This is a major advantage compared to other generation techniques, e.g. point-process models or alternating renewal approaches because the crucial step of the event separation can be omitted.

After the statistics are calculated for each observed location, they must be transferred to ungauged locations. In NiedSim2.x, all statistics are interpolated on a regular 1 km x 1 km grid using external drift kriging with the square root of the altitude as external drift (e.g. *Goovaerts, 2000*). The interpolated statistics enable to generate a precipitation time series for any arbitrary grid point.

All subsequent steps of NiedSim depend on the quality of the database and, thus, are vulnerable to measurement errors or inhomogeneities. Consequently, a proper analysis of the measured data is a very important part of NiedSim. This is why a significant part of my study concerns this topic. Nevertheless, NiedSim can only be as good as the measured input data.

4.4.2. 1 h time step generation

The generation of a precipitation time series in hourly resolution is subdivided into two parts, the setup of an initial time series (a) and its optimization (b). The optimization starts after the initial setup is finished such that both processes can be analyzed separately. This is an important aspect for the investigation of the performance later in my study. NiedSim works on an annual base, that is, each year of the desired time period is generated and disaggregated individually.

(a) Setup of initial time series In NiedSim2.x, the initial time series is generated by sampling from a distribution function of hourly precipitation values ϕ^* . The distribution function is a mixed function containing three different parts. A shifted gamma distribution describes values larger than 0.2 mm and an exponential distribution reflects values smaller 0.2 mm. Additionally, extreme values are drawn from the KOSTRA-DWD-2000 data set.

The actual generation of the initial time series is done as follows. At first, two values from the KOSTRA exponential distributed extreme values are drawn and one value from the KOSTRA Gumbel distribution. The following values are drawn from the gamma distribution until the annual volumetric percentage of precipitation above the threshold (0.2 mm) is reached. Finally, the time series is filled with small values from the exponential distribution until the total annual precipitation sum is obtained. All values are randomly distributed within the year. This procedure creates an initial time series with the correct precipitation amount and distribution, but an incorrect, random temporal structure.

(b) Optimization of initial time series The initial time series is optimized with a Simulated Annealing algorithm by minimizing the objective function

$$O = \sum_{i=1}^n \left(\frac{(X_{i,syn} - X_{i,tar})^2}{X_{i,ini}} \right)^{\omega_i} \implies \min \quad (4.1)$$

where X_i is a statistical property i of the precipitation time series and n is the total number of regarded properties. The subscript *syn* denotes the synthetic time series to be optimized, *tar* refers to the target, measured time series, and *ini* to the initial time series before the optimization. As the absolute values of each property can be highly different depending on the regarded statistic, the squared differences are standardized with respect to the initial series. The weighting factor ω_i can be used to prefer the optimization of certain statistics if their individual optimization is mutually exclusive.

In the Simulated Annealing optimization, two precipitation values are randomly swapped and after each swap the objective function is evaluated. If the difference is smaller after swapping than before, the change is accepted. If not, the swap might still be accepted with a certain probability to avoid local minima. The probability of acceptance of worse swaps is reduced during the optimization to converge towards the optimized hourly time

series. More information about the Simulated Annealing approach can be found in *Bárdossy* (1998).

In *NiedSim2.x*, the parameters ϕ and ψ of Tab. 4.1 are used as properties X_i . Theoretically, any kind of statistic could be used as target in the objective functions and makes the generation very flexible.

4.4.3. 5 min disaggregation

In the last part of *NiedSim*, the 1 h time series is disaggregated to a 5 min resolution with a realistic intra-hourly temporal structure, for which a similar Simulated Annealing algorithm is used as explained before. At first, the precipitation values are uniformly distributed between the twelve 5 min increments of one hour (block rainfall). In a second step, a small fraction is subtracted from one 5 min value and added to another 5 min value of the same hour. If the change is accepted or not, is evaluated using Simulated Annealing with an objective function comprising the 5 min statistics of Tab. 4.1.

4.5. Validation of synthetic precipitation

Two major questions arise when a precipitation model is validated. First, how well does the precipitation model works in general? Or in terms of *NiedSim*, is the Simulated Annealing optimization able to optimize the objective function? The second question is, how well does the synthetic precipitation represent the natural characteristics at the specified location?

The premier question is rather trivial for *NiedSim* because the target statistics obtained by measurements can directly be compared to the ones of the synthetic precipitation. The smaller the deviations, the better the optimization algorithm works.

The second question is a more difficult task for several reasons. Obviously, the synthetic precipitation would ideally represent the local characteristics of rainfall, if *all* statistics were *perfectly* fit. However, precipitation is highly complex with non-linear and high-dimensional dependences. For example, the autocorrelation can be calculated for different aggregations and time lags. Even if the scales from one minute up to one month with different time lags are evaluated, cross-dependences to many other statistics (e.g. seasons) exist. Therefore, one cannot assure that *all* statistics and dependences are reflected with a finite number of parameters. Additionally, there might be even statistics and relationships that are not considered at all because the process of precipitation is still not fully understood.

Another issue is, that the statistics cannot be *perfectly* fit because the true values of the population are unknown. All measurements can only reflect the statistics of samples which are influenced by measurement and sampling errors. Furthermore, many statistics describe the temporal global characteristics, but do not account for locally changing properties. An example is the autocorrelation, as we will see later, where different local temporal structures can exhibit the same autocorrelation. Consequently, a direct validation which verifies a perfect fit of synthetic and natural time series is impossible.

For the evaluation of simulations cross-validation techniques are often applied. Here, the model is calibrated by some parameters and the same parameters are then used for the validation, but for a different time period or location. Difficulties with the cross-validation occur for precipitation as mentioned by *Pui et al.* (2012). A cross-validation with different locations is not possible because long series of different, but close locations are usually not available. A splitting of the time series for calibration and validation period is also problematic, as long-term statistics might not be represented correctly by their short equivalents. *Pui et al.* (2012) proposes a kind of indirect validation, where statistics are validated that are similar to the ones used for calibration because they describe the same process, but were not explicitly used for calibration.

Another possibility of an indirect validation is the usage of a subsequent application. The target of modeling precipitation is actually not to reproduce data with nice-looking statistics, but to generate a time series that a subsequent application will treat as *correct* precipitation. *Correct* in this context means that the synthetic precipitation should not be distinguishable from observed data by the application, for example, a sewer system model. An advantage is that hereby not only the quality of precipitation simulations can be assessed. Additionally, it allows to infer from the model output back on the precipitation input, and thus it helps to decide which statistics are relevant for the application. Nevertheless, the direct evaluation of precipitation statistics are additionally necessary to verify that the results of the indirect validation are correct for the right reasons and do not just arbitrarily fit. In other words, the direct validation of precipitation statistics are necessary, but not sufficient for the evaluation of precipitation time series.

No standardized procedure exist for the direct validation of precipitation models in literature. Most authors use at least some global statistics like the autocorrelation (e.g. *Bárdossy*, 1998; *Onof et al.*, 2000; *Pui et al.*, 2012), p_0 (e.g. *Koutsoyiannis and Onof*, 2001; *Molnar and Burlando*, 2005; *Paschalis et al.*, 2014), or the cumulative distribution function (CDF) (e.g. *Verhoest et al.*, 1997; *Molnar and Burlando*, 2005; *Vandenberghe et al.*, 2011). The majority additionally assesses the extremes usually with IDF-relationships (e.g. *Koutsoyiannis and Onof*, 2001; *Gaume et al.*, 2007; *Pui et al.*, 2012). The evaluation of event-based statistics is regarded in particular, but not exclusively, for models using an alternating renewal approach (e.g. *Haberlandt*, 1998; *Salvadori and Michele*, 2006; *Müller and Haberlandt*, 2016). Furthermore, there exists some less common validation techniques applying, for example, spectral analysis (*Paschalis et al.*, 2014), copulas (*Salvadori and Michele*, 2006; *Vandenberghe et al.*, 2011) or mass curves (*Verhoest et al.*, 1997).

An indirect validation is not very common and only used by a few authors for sewer models (*Haberlandt*, 1998; *Hingray and Haha*, 2005; *Gaume et al.*, 2007; *Andrés-Doménech et al.*, 2010; *Müller and Haberlandt*, 2016).

The reason that many different validation techniques are applied can be explained by the variety of models and their diverse input parameters as well as by the complexity of rainfall. Nevertheless, the objective of a proper validation should be to find a way which is comparable, fair, consistent, independent of the generation method, and finally, appropriate for respective applications. This is why I will pay much attention to define such a (indirect) validation method in the present study.

I would like to give a general statement at the end of this chapter. Observations always reflect the characteristics of the past. Usually, one assumes that information of the past includes indicators about the future, and thus the data can be used to predict the future. However, this assumption need not to be true and, obviously, cannot be proved. Consequently, even if measured time series were perfectly represented by synthetic data, the NiedSim simulations will not necessarily reflect future observations.

5. Sewer systems

Sewer systems are one of the oldest parts of the urban infrastructure and essential for the human wellbeing. Sewers serve two major purposes: First, the hygienic collection and disposal of wastewater, and second, the prevention of flooding by safely transporting and retaining stormwater. Over the years, these traditional and emission based objectives have been changed towards immission based approaches including the ecosystem river. In particular, the Water Framework Directive of the European Union in the year 2000 (EU, 2000) changed minds with its tasks to achieve a good ecological and chemical status of all surface waters. Nowadays, both objectives should be achieved in conjunction with an ecological, chemical and quantitative protection of water bodies as well as the security for further usage like fishing, swimming or drinking water supply (DWA-A100, 2006).

The most significant changes concern the handling of stormwater. Stormwater is termed as surface runoff from urban areas caused by precipitation. The expansion of impervious areas has changed the natural water cycle of urban areas and causes increased surface runoff volume, larger peak flows and higher flow variability, along with decreased infiltration processes and reduced lag times (Jacobson, 2011). All effects have a negative impact on flooding problems (Hammond *et al.*, 2015) as well as on combined sewer overflows (CSOs) (Semadeni-Davies *et al.*, 2008). In the past, the target was to collect and discharge stormwater as much and as fast as possible for which large sewer pipes were necessary. However, the larger the sewers are, the higher the installation and maintenance costs will be. Therefore, these days the focus is changing towards stormwater source control approaches, where the goal is the reduction of storm water discharge to a minimum by local retention, infiltration and treatment methods (DWA-M153, 2007). These integral urban water management approaches are also known as low impact development (LID), sustainable urban drainage systems (SUDS) or water sensitive urban design (WSUD) and include structures like grass swales, green roofs or permeable pavers. They all share the same target of a decreased discharge leading to less hydraulic pressure in sewer systems, to a better treatment of waste water in the wastewater treatment plants (WWTPs), and finally, to less pollution of the receiving waters bodies. The current trend is towards LID approaches as shown in Elliott and Trowsdale (2007) or in the projected DWA-A102 (2016). Nevertheless, in most parts of the world, even in developed countries, classic sewers are still the primary system for the collection and discharge of precipitation. Therefore, the focus of this work is on the design and simulation of classic sewers without explicitly focusing on LIDs.

5.1. Overview of sewer systems

Before details about sewers and their relation to precipitation can be explained, some technical terms have to be defined. In a sanitary sewer, the sewage from domestic, commercial and industry is collected and transported to the WWTP. The conveyed flow is called wastewater. Water of drainages or groundwater that infiltrates into the sewer unintentionally is named foreign water. Wastewater in addition with foreign water is referred to as dry weather flow. If stormwater and wastewater are discharged in individual pipes, the sewer is called a separate system. If both are collected together, the sewer is referred to as a combined sewer.

In a separate sewer, the polluted wastewater is conveyed completely to the WWTP in a foul sewer. The stormwater, collected in a storm sewer, is assumed to be less polluted, and thus can be directly discharged into the receiving water without passing the WWTP. However, the assumption of stormwater being less polluted has been disproved during the last decades (e.g. Göbel *et al.*, 2007), which is an additional reason for the increasing trend towards LID techniques.

In combined sewer systems, wastewater and storm water are conveyed together as combined sewage discharge. The pipes must be larger as compared to the foul sewer due to the stormwater discharge, but cannot be designed on the maximum occurring flow condition for the following reason. Large flows caused by intensive precipitation events occur only rarely, and consequently most of the time the pipes as well as the WWTP would be oversized. As a result several problems would arise like increased deposition of sediments in the pipes during dry periods, problems with the microbiology in the WWTP and exorbitant costs for the construction and the maintenance of the sewer and the WWTP. Therefore, in a combined sewer system the discharge Q_d to the WWTP has to be limited without losing its general objectives: the collection and discharging of sewage without negative impacts on the ecosystem river. In Germany, the limit of Q_d , which can usually be treated in the WWTP, is twice Q_m plus the flow of the sewer infiltration water (ATV-A128, 1992).

The limitation of the discharge to the WWTP can be achieved either by a detention and delay of the flow with additional storage capacity in the sewer, or by a diversion of the flow. In the latter case, only a fraction of the flow is conveyed to the WWTP and the rest is directly discharged into the water bodies. This task is accomplished by CSO structures. These flow diversion structures contain a throttle where a maximum Q_d (continuation flow) is defined, and a weir that separates the exceeding flow (spill flow) from the continuation flow. If the maximum continuation flow is exceeded, CSO events occur where the spill flow is discharged untreated into a receiving water body. As the pollutant loads during these events can be very high, the focus of the design of combined sewers is on minimizing the ecological impacts of CSO events.

From an engineering point of view, sewer systems contain three classical fields (Giesecke and Haberlandt, 1998): hydraulic design of sewer networks (Chap.: 5.1.1), planning of combined sewer systems (Chap.: 5.1.2) and sewer system management (Chap.: 5.1.3). Although precipitation is the most important variable for all three tasks, there are major differences related to the relevance of certain characteristics of precipitation.

5.1.1. Hydraulic design of sewer networks

The objective of hydraulic design is to reduce system failures of sewer networks. Critical design flow rates are defined to avoid backwater in the pipes, surcharging of manholes, and finally flooding of surface areas or cellars. Consequently, the major interest in the hydraulic design is on the peak flow rates (*Schilling, 1991*). The design flow rates depend on the type of sewer system.

A foul sewer is mainly designed on the maximum occurring dry weather flow. Although the dry weather flow of a single house lateral can be very intermittent, it is relatively continuous at the outlet of a foul sewer (*Butler and Davies, 2011*). Hence, the maximum dry weather flow can be estimated rather well.

The pipes of storm sewers and combined sewers must in general be larger compared to foul sewers because the stormwater flow can exceed several times the dry weather flow during heavy precipitation events. Thus, the critical design rate depends mainly on the maximum stormwater discharge. The maximum annual storm water discharge can highly vary from one year to the next depending on the occurrence of large or extreme precipitation events. Therefore, the maximum annual flow of a certain year, or even within a given time period, cannot be predicted exactly. It can only be described statistically by an estimation of its probability of occurrence like by IDF-curves explained for precipitation in Chap. 2.2.3.

However, the return period of a precipitation event is not equivalent to the return period of a flooding event; the latter can be much higher. The reason is that a sewer consists of more components than just the pipes. For example, gullies, manholes, and the fact that pipes are installed at least one meter below the surface increase the storage compared to the pipe-full capacity on which the design criteria are usually based (*Butler and Davies, 2011*). According to the European Standard DIN EN 752, the criteria for flood protection is the flooding frequency of an area or property. However, the actual extension of a flooding event and its respective frequency is difficult to represent. Thus, in some countries like in Germany, the primary design criteria is the exceedance flow with respect to the ground level that emerges if flow occurs on the surface of the sewer, or if water cannot enter the sewer. Depending on the vulnerability of the area, the exceedance flow frequency must be less than once in one to ten years, or accordingly, must have a return period of at least ten (to one) years (*DWA-A118, 2006*). Only if a exceedance flow at some locations of the sewer is demonstrated in the simulation, a more detailed analysis have to be conducted including the proof of flood protection of certain properties or vulnerable places.

The stormwater flow that is necessary to evaluate the occurrence of surface flooding can be obtained by artificial design storms based on IDF-curves of precipitation data (*DWA-A118, 2006*). An alternative to design storms are observed time series of intense rainfall. Here, continuous, measured series are separated first into single events. Afterwards, certain events are selected to be simulated that are supposed to cause surface flooding.

In conclusion, only large and extreme precipitation intensities are crucial for the hydraulic design of sewers. The respective precipitation events can be described sufficiently with the concept of IDF-curves and design storms.

5.1.2. Design and simulation of combined sewer systems

In combined sewer systems, not only the pipes of the sewer have to be designed, but also the CSO structures. In Germany, an emission based approach is applied traditionally for this purpose. That is, the integral pollution emissions caused by the effluent of the WWTP as well as by CSO events are considered. The most widely used guideline for design and simulation of combined sewer networks in Germany is *ATV-A128* (1992). Here, the ecological conditions of the individual receiving water bodies are not yet regarded. Immission based concepts are described in *BWK-M3* (2007) in cooperation with the projected *DWA-A102* (2016), a subsequent guideline of *ATV-A128* (1992).

In the emission based approach of *ATV-A128* (1992), the chemical oxygen demand (COD) is taken as a general measure of the pollution. The criteria for design and verification of the performance of a combined sewer system are defined with a mean annual COD load of the receiving water bodies under average conditions caused by storm water. The relevance of mean annual values are in contrast to the hydraulic design, where the design criteria are based on maximum values. A mean reference case for the COD emission is defined, where the mean annual precipitation depth (800 mm) is stated as well as the relation of COD concentrations between dry weather flow, storm water flow and effluent of the WWTP (600 : 107 : 70). By using the reference case and applying a mass balance of the COD loads, a theoretical, maximum acceptable CSO rate can be calculated. It is defined as the maximum annual CSO volume divided by the annual runoff.

The actual design of the combined sewer systems contains three different steps. In a first step, the minimum total storage volume of the combined sewer system is obtained by using the predefined reference case explained above. As the actual conditions of waste water and storm water can highly vary for different areas, the mean reference case is adjusted to the regional conditions. The necessary storage volume is then defined by two main variables depending on the local situation. The first parameter is the acceptable CSO rate, which is related to the actual pollution of the CSOs. That is, the more the spill flow is contaminated, the more storage volume must be provided. The second variable is the run-off discharge rate q_r that depends on the capacity of the biological treatment of the WWTP. q_r is defined as the rainfall run-off per impermeable, connected catchment area. The larger the capacity of the WWTP, the higher q_r can be and the less volume is necessary. According to *ATV-A128* (1992) the standard design procedure of determining the necessary volume is valid for q_r up to $2.0 \text{ l}/(\text{s} \cdot \text{ha})$ as well as for specific storage volumes V_s between 5 and $40 \text{ m}^3/\text{ha}$. V_s is specified as storage volume per impermeable, connected catchment area¹. V_s of $40 \text{ m}^3/\text{ha}$ is supposed to be the limit due to economical reasons. A q_r larger than $2.0 \text{ l}/(\text{s} \cdot \text{ha})$ can be possible if the WWTP has a substantially larger capacity compared to the design dry weather flow. In these cases, the performance has to be proven via sewer simulations.

After the required storage volume is defined, it can be rearranged among several CSO structures in the sewer system. The criteria of the allocation is such that the COD emission will not be increased compared to the undistributed case, where all storage volume is

¹The specific area of V_s and q_r is always defined as impermeable, connected catchment area throughout this study.

installed just in front of the WWTP. For small and ordinary sewer systems, a simplified distribution approach can be applied, otherwise the criteria must be proven with a long-term combined sewer simulation. In a final step, each single CSO structure is individually designed.

The CSO structures can be separated into overflow structures with and without storage capacity. CSO structures without storage consist in their simplest form of a weir which only diverts the flow. The objective of these structures is to cut the peak of large combined sewage flow intensities. As the spill flow is not treated at all and the polluting load might be large due to first flush effects, the run-off discharge rate q_r must be large, that is, at least $7.5 - 15 \text{ l}/(\text{s} \cdot \text{ha})$ (ATV-A128, 1992). The first flush effect denotes the increased pollution load at the beginning of a rainfall event after a long dry period (Barco *et al.*, 2008). Although, the threat of contaminants for such CSOs is rather low due to the high dilution, these structures should be installed where the combined sewage flow is the least polluted.

CSO with storage can be further categorized by type (bypass tank, pass-through tank, a combination of both or sewers with in-pipe storage capacity) and by arrangement (on-line, off-line or mixed design) (ATV-A166, 2013). I do not give detailed information about advantages, disadvantages or the design in this study, as my focus is not on the design or pollution reduction aspects of storage structures. An important point with respect to my study is the emptying time t_e of the structures calculated as V_s divided by q_r . Some combinations of V_s and q_r are less important in practice because t_e should not exceed 10 - 15 hours (ATV-A128, 1992).

As explained before, the critical design rate of combined sewer systems is the capacity of the WWTP, or in terms of flow, the discharge rate q_r to the WWTP. However, the stormwater runoff can already exceed q_r even for medium precipitation events and lead to CSO events several times a year. Therefore, the exclusive consideration of extreme precipitation events is not sufficient, which is a major difference to the hydraulic design of sewers. Furthermore, the qualitative and quantitative evaluation of CSO events highly depends on the pre-filling and thus, on the antecedent precipitation events. Even rather small precipitation events can cause overflow if they occur after a wet period where the storage tanks have already been filled. On the other hand, larger events may be completely buffered in the system if enough storage capacity is available after a dry period. The antecedent precipitation history is also important if pollution emissions are considered. The longer the dry weather period, the more the contaminants in the deposits can accumulate.

In conclusion, for the design of combined sewer systems, the complete precipitation time series can be relevant and must reflect the local, average long-term characteristics. The necessary, but rarely available continuous, long-term precipitation time series for the simulation of combined sewers are the major motivation of the generation of synthetic time series of NiedSim.

5.1.3. Sewer system management

The target of sewer system management is to improve the objectives of both previous tasks (the hydraulic design and the design and simulation of combined sewers) by active inter-

ventions in the system via flow regulations. The monitoring and control of the sewer system in real time is also known as real-time control (RCT). That is, the transport, storage and treatment facilities of the sewer systems are optimized based on real-time information (García *et al.*, 2015). Several objectives can be concerned by applying real-time control (RCT) like the prevention of flooding, the reduction of overflows or the optimum use of existing storage capacity (Schütze *et al.*, 2011). In all cases, the precipitation input requirements must be most detailed compared to the previous applications because even single precipitation events should be reflected correctly.

These applications are not considered in my study because they are the most complex ones. Before they can be investigated, the synthetic precipitation should be evaluated for the previous tasks. Without having a time series representing the extremes and the mean correctly, it is useless to validate their applicability for tasks requiring an even more detailed temporal structure.

5.2. Sewer model representation

A sewer model mimics all surface and subsurface processes from precipitation input to the final flow to WWTP or the spill of CSOs into the receiving waters. If, additionally, a detailed model of the WWTP, the river system, or other social and ecological components are included, it is referred to as integrated sewer system modeling (Bach *et al.*, 2014).

A classical sewer model can be separated into storm water quality and quantity modeling, where the latter one includes four major parts. The first part is the hydrological characterization of the runoff generation, where the amount of precipitation is transformed into runoff volume. It is followed by the representation of the overland flow that specifies the temporal progress of the runoff between its initial place and the inlet of the sewer system as well as the combined runoff from several subcatchments. The third part characterizes the flow processes inside the sewer system including important components like pipes, pumps or CSO structures. The last component describes the surface flooding, that is, if the sewer is fully surcharged, and thus flow leaves the sewer system or cannot enter. The quantity of flow in the sewer is described as water flow [m^3/s] or, additionally, in more complex hydrodynamic models as water depth [m].

The discharge quality can be simulated statistically by regression models or empirically by buildup and washoff functions. The latter ones describe the sedimentation of pollutants in the channel bed of the sewer during dry periods and the erosion of the pollutants during wet periods. The transport and transformation of pollutants within the sewer can be described via five different processes: chemical reactions, biological, ecological and physiochemical processes (adsorption, desorption, absorption) as well as physical transportation (advection and diffusion) (Zoppou, 2001).

The modeling of the quantity is the driving variable for the hydraulic design tasks of sewer systems. The quality parameters are important to characterize CSOs, and thus are relevant for design and simulation of combined sewer systems as well as for sewer system

management. However, even if the quality parameters are of major interest for these tasks, the quantity cannot be neglected because of the interdependence between quality processes and quantities *Zoppou* (2001). Actually, CSO statistics about the quantity, like the overflow frequency or volume, can already be used to infer about the quality parameters *Lau et al.* (2002).

There are some additional arguments why I avoid the quality modeling in this study. First, I use sewer simulations as an indirect method to validate the synthetic precipitation data. The precipitation input is directly related to the quantity of runoff and discharge, but not to quality parameters. Second, single processes of the wastewater quality modeling are highly complex (*Ashley et al.*, 1999) and may never be fully understood. Many different approaches exist for quality modeling based on empirical relationships that involve high uncertainty. In contrast, the hydraulic behavior of the discharge quantity can be described physically using hydrodynamical descriptions. Third, even for most quality models, the quantity is a relevant and necessary part. If the quantity is not represented well, the informative value of the quality modeling is useless.

The quantity parameters of the flow in the sewer system are obtained from highly complex natural processes. Some processes, for example the overland flow, are very complicated such that only empirical relationships can be established, while others can be fully described physically, but are computationally very expensive. Therefore, simplifications of the processes have to be applied depending on the desired objective. A crucial point is to balance the simplification of the processes and the accuracy of the model output that is necessary for a specific task. Therefore, models with various complexities are used depending on the purpose for which the required accuracy of the input data differs. The more detailed a model is, the better the quality of the input data must be.

Sewer models can be classified according to their temporal and spatial characteristics. The spatial resolution of a model can either be distributed or lumped. In a lumped model, the characteristics of the entire catchment are represented on average, whereas a distributed model can take different subcatchments into account. The temporal resolution of the model can either be continuous or event-based, where only certain events are simulated. The advantage of the event-based method is, that processes of individual events can be modeled very detailed without requiring enormous computing power. For design tasks usually detailed models are set up for which only single events are simulated. Whereas for simulation and optimization of combined sewers, rather simple models are developed, but long-term simulations are performed.

Classic sewer models are built as deterministic models where one set of input variables results in one specific set of output variables. Therefore, the uncertainty introduced by the natural random process is not explicitly regarded in sewer models. However, it can be considered, for example, with Monte Carlo methods, where the model runs many times with altered input parameters (*Aronica et al.*, 2005).

In the following sub-sections, the four model parts characterizing the quantity of the flow are explained in more detail. These parts are relevant for the setup of the synthetic sewer for the indirect validation later in my study.

5.2.1. Runoff generation

Precipitation falling on a surface area is the driving variable of runoff. However, several effects influence the fraction of precipitation that is transformed into actual surface runoff. The amount of precipitation that contributes to runoff is referred to as effective rainfall and depends on depth and duration of a precipitation event. The higher the depth or the longer the duration, the larger the effective rainfall is. Therefore, the crucial factor of runoff generation is the description of the catchments surface characteristics. These include, for example, the vegetation, the soil type, the moisture content, the surface slope, and in particular, the fraction of impervious area. The larger the percentage of impervious area, the higher the effective rainfall is. If the contribution of pervious area to urban runoff is significant, is not yet completely understood (*Fletcher et al.*, 2013) and is usually neglected.

The losses reducing the effective rainfall can be separated into initial losses and continuing losses. Initial losses include the wetting of surfaces and interception losses of vegetation. Both effects reduce the precipitation only once at the beginning of each event and their reduction range between 0.3 mm to 0.7 mm (*ATV-DVWK-M165*, 2004). As they are rather small, they are often disregarded or combined with the depression storage (*Butler and Davies*, 2011). The depression storage is also part of the initial losses, and reflects the storage capacity of small depressions on the surface. Typical values of the depression storage for impervious areas are in the range of 0.5 mm to 2.0 mm (*ATV-DVWK-M165*, 2004).

Continuing losses like evapotranspiration or infiltration decrease the effective rainfall throughout the entire event. Evapotranspiration is necessary to close the water balance, and thus is important for applications where the whole urban water cycle is regarded, for example, for stormwater retention and infiltration strategies (*Fletcher et al.*, 2013). However, as evapotranspiration is only 2 to 3 mm per day, it can usually be neglected in applications where high intensities are most critical such as hydraulic design (*Butler and Davies*, 2011). On the other hand, if continuous sewer simulations are performed, evaporation is a necessary process for recovering the depression storage. The actual evaporation depends among others on humidity, wind speed and temperature. For sewer simulations, however, the complex process is usually simplified to mean daily evaporation values reflecting an annual cycle. The infiltration losses can be described using, for example, the empirical Horton's equation. This equation defines the infiltration rate per unit time depending on several factors describing the soil characteristic. As only impervious areas are considered in this study, the infiltration process is not further described here.

A simplified alternative is the combination of the initial and continuing losses in a runoff coefficient approach. Here the effective rainfall rate is proportional to the net precipitation rate times a factor. The factor called runoff coefficient can either be a simple time invariant constant or a variable depending on different precipitation events. Typical runoff coefficients vary between 0.70 to 0.95 for impervious areas and can be below 0.30 for pervious areas (*Butler and Davies*, 2011). Depending on the application, mean or peak runoff coefficients are used. The runoff coefficient approach is often applied in combination with the classical rational method for hydraulic design. Then, a peak runoff coefficient is calculated depending on the slope, percentage of impervious area and rainfall intensity as described, for example, in *DWA-A118* (2006).

Another widely applied method for calculation of effective rainfall is the curve number approach developed by the United States Department of Agriculture (USDA) formerly Soil Conservation Service (SCS), where the runoff is empirically related to precipitation and surface characteristics (NRSC, 2004).

5.2.2. Overland flow

Once the effective rainfall is obtained, it can be related to the respective area resulting in surface runoff, also known as overland flow. This section is about the representation of the relation between the hyetograph of the effective rainfall and the hydrograph of the overland flow. The process reflects the transformation of the areal-temporal precipitation data over each subcatchment into point-temporal information at the lowest point of the subcatchments, as well as the superposition of the hydrographs from several subcatchments at the final inlet of the sewer system. An exact physical description of the overland flow is very difficult due to the high heterogeneity of different surfaces. The surface runoff is usually approximated either physically by the continuity equation and a simplified equation of motion, or hydrologically with the unit hydrograph concept.

Even if a physical model is used, some simplifications must be applied to the full equations of motion. A method commonly used is based on the continuity equation describing the mass balance between effective rainfall and runoff. The runoff is then approximated with the empirical Manning equation. This equation relates the flow to the hydraulic radius approximated as water depth, the surface slope and the Manning coefficient. The Manning coefficient describes the roughness of the surface and is an empirical parameter that ranges from 0.01 for smooth asphalt up to 0.12 or larger for vegetated surfaces (Rossman and Huber, 2016).

An alternative is a hydrological lumped model where a unit hydrograph is utilized that combines all specific surface characteristics in a few parameters including translation and attenuation of runoff. The unit hydrograph relates the unique response runoff curve of a catchment to a unit depth of precipitation within a unit duration. It is based on two major assumptions, first, the unit hydrograph is time-invariant and second, several input signals can be superimposed to one output signal (Dooge, 1959). Under these assumptions, the effective rainfall can be split into fractions with unit duration, for which the runoff response can be individually obtained. The total runoff curve of the precipitation event can then be calculated by the superimposition of all individual runoff responses. The unit hydrograph can be calibrated either by synthetic hydrographs for example the standard unit hydrograph (ATV-DVWK-M165, 2004), by a cascade of linear reservoirs (Beven, 2012) or by time-area-diagrams (Butler and Davies, 2011).

5.2.3. Sewer flow

The pipe and channel flow inside a sewer are well known processes and can be fully described using physical equations. As sewer and overland flow are two decoupled systems

where the hydrograph of overland flow is not influenced by sewer flow conditions, the sewer can be assessed without acknowledging for uncertainties of the surface flow. The flow can be calculated with the shallow water wave equations (Saint-Venant equations) consisting of a continuity and a momentum equation (Zoppou, 2001). The solution of the full Saint-Venant equations yields flow rates as well as water depths. The dynamic equation incorporates all terms that are necessary to describe all parts of the hydrodynamic processes (wave translation, wave attenuation and flow acceleration). However, the complete solution of the equations is computationally expensive, and thus several simplifications can be applied. If steady state conditions are assumed and if convective and local acceleration terms are neglected, the dynamic wave representation can be simplified to the diffusive wave approximation. The diffusive wave can still account for wave translation and diffusion, and thus is able to describe effects like backwater and flow reversal. If, additionally, the flow is uniform, the wave attenuation term (the pressure slope) can be disregarded. In the resulting kinematic wave approximation, the wave is translated with a kinematic wave speed through the system without changing its shape. The flow can vary temporally and spatially, but the kinematic wave can neither account for backwater effects nor for channel storage.

In contrast to the full or approximated Saint-Venant equations, several alternatives based on empirical approaches exist. The simplest one reflects only the translation of the flow hydrographs from upstream to downstream depending on the travel time. These models may not be applied for large and flat sewer systems because only for small and steep sewers the channel storage can be neglected (ATV-DVWK-M165, 2004). More sophisticated approaches are able to reflect the attenuation effect and are based on hydrological models. These models use the continuity equation of the Saint-Venant equations, but approximate the momentum equation with empirical relationships. The parameters can either be obtained using a cascade of reservoirs based on geometric and hydraulic considerations for example with the Kalinin-Miljukov approach, or from in- and outflow relationships like in the Muskingum approach. According to ATV-DVWK-M165 (2004) the first one is usually applied for sewer systems. However, all kinds of hydrological models cannot account for dynamic effects like backwater, channel storage, pressurized flow or flow reversal, and thus they should only be applied if these processes can be neglected. Therefore, typical applications for hydrological models are long-term planning simulations of combined sewer systems.

The first modern and simplest model approach is the rational method. According to Burian and Edwards (2002) it has already been invented in the nineteenth century, but is still in use (Young *et al.*, 2009). In the rational method the peak flow rate is assumed to be proportional to the precipitation intensity linked by an area and a rational runoff coefficient. As large precipitation intensities can be described by a design storm (Chap. 2.2.3), the design runoff can directly be derived from the rational formula. The design storm should cover at least the worst case scenario, which occurs when the entire area contributes to the runoff at the outlet of the sewer. Therefore, the duration of the design storm must be at least the time of concentration (Butler and Davies, 2011). The time of concentration is the necessary time of runoff flow to reach the outlet of the sewer from the furthest point of the surface of the catchment. Knowing the concentration time and the return period of the design storm one can derive the design precipitation intensity from IDF-curves. The design runoff can then be calculated using the rational formula if a certain runoff coefficient of the catchment area

is assumed. By applying the rational method only peak flows can be obtained. Therefore, it is appropriate for hydraulic design purposes, but only recommended for small and simple sewers (*DWA-A118*, 2006).

5.2.4. Surface flooding

In cases where the capacity of the sewer is exceeded, so called exceedance flow occurs on the surface and can cause flood events. Several different approaches exist to model exceedance flow or even surface flooding. These models can be very simple, for example, by extending the sewer model with additional storage capacity on the surface, or by including simple flow routing approaches according to the surface topography (*Butler and Davies*, 2011). On the other hand, highly complex coupled 1D sewer and 2D surface models can be applied, where both are based on the full or approximated Saint-Venant equations (e.g. *Schmitt et al.*, 2004). However, as surface flooding is not part of my investigation, I will not describe these models in more detail.

6. Setup of an artificial sewer model for an indirect validation

A major aspect of the validation of precipitation time series in my study is their usability for different tasks in sewer system modeling (Chap. 4.5). Depending on the task, certain precipitation statistics are more relevant than others. Consequently, one should clearly define the applications for which the precipitation time series are intended to be used. In my study, I focus on the hydraulic design of sewers as well as on the design and simulation of combined sewers.

Even if the applications are restricted to the ones stated above, it is still challenging to define a general procedure for an indirect validation. As one cannot test the precipitation time series with all existing sewers in the real world, one needs to find or define one or more representative sewers. However, most of the sewers are unique in their characteristics based on the local boundary conditions. A case study of a small system might not reflect the broad range of existing sewers, and would just show the characteristics of a very specific system. On the other hand, if a large system was applied, the complexity and interaction of different structures would increase. Therefore, it would be very difficult to infer from the results of a complex sewer simulation back on the input precipitation characteristics. Clear inferences between input and output are a necessary attribute of a sewer if it is supposed to be used not only for validation, but also as a tool enabling an iterative improvement of the generation of synthetic precipitation. Consequently, an artificial sewer system should be applied in such cases.

However, only for the second task (design and simulation of combined sewer systems) an artificial sewer was set up. The first task (hydraulic design) is very difficult to be investigated with synthetic sewers because of the complex interactions of drains, structures, manholes, pumps and single conduits that differ in diameter, slope and roughness. All these aspects impact the surcharging and flooding behavior and, thus, influence the final design. Therefore, simulations of individual conduits or simple systems are likely to be not representative for real case sewers.

For the design and simulation of combined sewer systems, it is much easier to find a manageable number of representative cases because only annual averages are relevant. Thus, the representation of the sewer needs not to be as detailed and can be modeled hydrologically.

In a previous evaluation (Müller *et al.*, 2014), I applied the artificial sewer system described in *ATV-A128* (1992) for an indirect validation. The results were not satisfying because only four different structures were defined in the artificial system. I introduced several modifications of the system by *ATV-A128* (1992) regarding the specific volume V_s and the specific

discharge q_s to increase the number of configurations. However, as the CSO structures were not independent of each other, the effects of precipitation input and sewer output could not be clearly related to the different configurations of the system.

In a second approach, I applied the artificial sewer system described in *Drechsel* (1991). According to him the V_s as well as the q_r are the most important parameters and are varied for different CSO structures. He stated that other parameters like the flow time, the dry weather flow, the geometry of the conduits or the extend of the catchment area cannot be neglected for calculating absolute values, but are less relevant for a relative comparison of precipitation time series.

Initially, I applied both setups in the hydrological combined sewer model KOSIM developed by itwh (*itwh*, 2009). In general, the results of the indirect validation were satisfying using the setup of *Drechsel* (1991). However, some questions with respect to the parameters as well as to the model setup remained unclear. In the course of modifying the setup of the artificial sewer used for this study, I changed the applied model to the hydrodynamic model SWMM developed by the US Environmental Protection Agency (EPA) (*Rossman and Huber*, 2016) because of its higher flexibility. In the following, I will explain the configuration of the original model of *Drechsel* (1991) and will give some general interpretations. Furthermore, I will illustrate my modifications of his artificial sewer and I will show the final implementation in SWMM.

6.1. Configuration and interpretation of the artificial sewer

The original system of *Drechsel* (1991) is separated in five different blocks with varying V_s $\{0, 15, 30, 50 \text{ m}^3/\text{ha}\}$. Within each block q_r is altered between $\{0.3, 0.5, 1.0, 2.0 \text{ l}/(\text{s}\cdot\text{ha})\}$. The combination of all parameters should cover a large bandwidth of existing stormwater tanks with overflow. An additional block complements the system, representing the CSO structures without storage. Here, q_r is much larger and set to $\{5, 10, 15, 30 \text{ l}/(\text{s}\cdot\text{ha})\}$. The different configurations are summarized in Tab. 6.1. As the dry weather flow is neglected, the throttle discharge q_d of each structure is equal to q_r . All CSO structures are connected to catchment areas of the same size and runoff characteristics. They discharge individually to the WWTP or receiving water, and, thus, can be analyzed independently.

The structures with different V_s and q_r of the artificial system respond like different filters that are applied on the precipitation time series. The larger the specific volume or discharge, the less precipitation spells will lead to overflow. An increase of q_r can be regarded as a filter cutting low intensity rainfall spells as the flow caused by such spells can directly pass the throttle. That is, the larger q_r is, the more intense (peaked) the flow must be to induce CSOs. Flows with a high intensity are only caused by heavy precipitation spells that are usually short in Germany. Therefore, the structures RR1 – RR4 rather show the behavior of intense convective rainfall events with an increasing focus on the extremes. From an engineering point of view, these structures represent simple weirs with its objective to cut the peak flows during heavy rainfall events.

	$V_s [m^3/ha]$	$q_r [l/(s \cdot ha)]$	$t_e [h]$		V_s	q_r	t_e		V_s	q_r	t_e
B01⁻	0	0.3	0.00	B31	30	0.3	27.78	R1	-	5.0	-
B02⁻	0	0.5	0.00	B32	30	0.5	16.67	R2	-	10.0	-
B03⁻	0	1.0	0.00	B33	30	1.0	8.33	R3	-	15.0	-
B04⁻	0	2.0	0.00	B34	30	2.0	4.17	R4	-	30.0	-
				B35⁺	30	5.0	1.67	R5⁺	-	50.0	-
B11	15	0.3	13.89	B51	50	0.3	46.30	R6⁺	-	75.0	-
B12	15	0.5	8.33	B52	50	0.5	27.78				
B13	15	1.0	4.17	B53	50	1.0	13.89				
B14	15	2.0	2.08	B54	50	2.0	6.94				
B15⁺	15	5.0	0.83	B55⁺	50	5.0	2.78				

Table 6.1.: Overview of the characteristics of CSO structures including the specific storage volume V_s , the run-off discharge rate q_r and the calculated emptying time t_e . Structures marked with ⁻ are included in the original system of *Drechsel* (1991), but not in the current system. Additional structures included in this study are marked with ⁺.

The increase of V_s shows the behavior of an integral filter where rainfall spells up to a certain depth do not lead to CSOs. In order to cause CSOs, a specific total precipitation spell depth must be exceeded which is related to the storage volume. The response of structures with storage is less dependent on the precipitation intensity. Therefore, the structures B51 – B54 reflect all events having a large total volume, that is, heavy and short storms as well as long lasting medium events. The less V_s (from B34 to B01), the less the precipitation time series is filtered and the more the medium and small spells have an influence on the CSO characteristics.

The combination of different V_s and q_r leads to different theoretical emptying times t_e . The higher t_e , the more likely it is that a partially filled storage affects the characteristics of a precipitation spell. One can also describe t_e as the memory of the storage structure. The emptying time is important for an event-based analysis because it defines the minimum time between two precipitation events such that they can be assumed to be independent. The emptying time also shows the relevance of the structures in practice because according to *ATV-A128* (1992) it should not be larger than 10 – 15 hours.

I would like to mention that the indirect validation responds like a low pass filter, where the small flows induced by small precipitation spells do not influence the overflow capacity. A single structure starts to cut from smaller towards larger precipitation spells (either in intensity or in total depth). One has to keep in mind that larger events are always included in structures of the same, but less filtered group. These structures show the additional, but not exclusive influence of smaller events. That is, if large spells are not represented correctly, the effect can be seen in all CSO structures that are less filtered. For example, if extreme intensity spells are too small, and thus RR4 is underestimated, the structures RR1 – RR3 (and even others if the total volume is large enough) will be underestimated as well. If they were not, the interpretation is not that less extreme events fit well, but rather that smaller

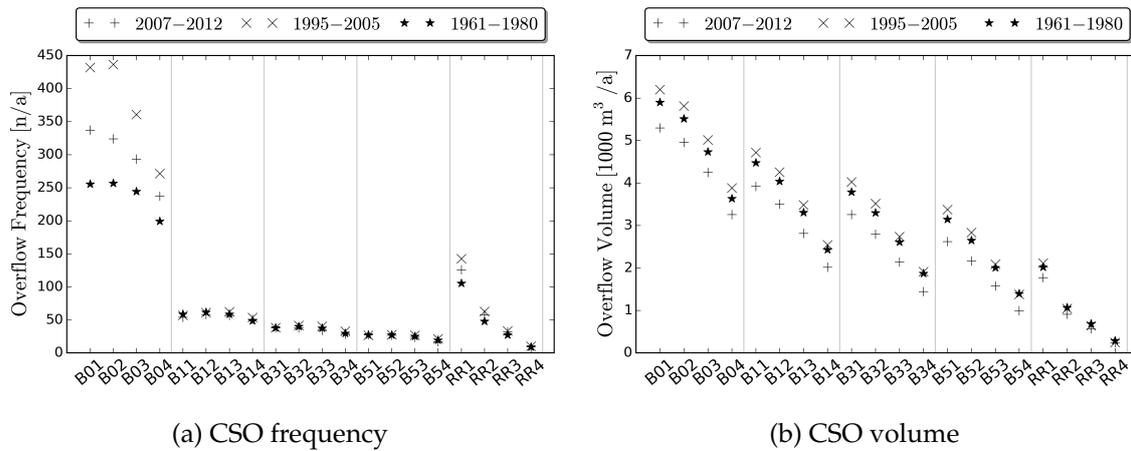


Figure 6.1.: Mean annual CSO frequency and volume of Freiburg for three time periods simulated with the original artificial sewer of *Drechsel* (1991) and the model KOSIM.

events must be overestimated in order to compensate for the underestimation of extreme events.

6.2. Modifications of the original artificial sewer

In this study, I applied the following adjustments to the original artificial sewer of *Drechsel* (1991). First, the structures of the first block (01 – 04) are not included anymore. These structures have a rather small q_r , but do not contain any storage, and thus are not realistic cases. One could argue, that for an indirect validation the entire structure should be represented correctly including even very small precipitation values. Such spells might be important in more complex sewer systems due to a complex interaction between different CSO structures. However, these structures are highly sensitive such that even different precipitation measurements highly influence the CSO behavior.

In Fig. 6.1a, the impact of three time periods representing different measurement devices on the CSO frequency is clearly visible. As explained in Chap. 3, the structure of small precipitation values depends on the measurement device. The time series of the period of digitized paper charts (1961 – 1990) is much smoother compared to the period 2007 – 2012, which leads to a systematic increase of the CSO frequency of very sensitive structures (B01 – B04). In contrast, the CSO volume (Fig. 6.1b) and duration (not shown) of both time periods are rather similar. Consequently, the deviations can be supposed to be artifacts caused by the too sensitive structures B01 – B04. In order to have a more robust and realistic indirect validation with respect to real sewers, I do not use these structures in the modified sewer.

According to *ATV-A128* (1992) the minimum q_r for structures without storage should be 7.5 – 15 $\text{l}/(\text{s}\cdot\text{ha})$. R1 and partially R2 are below this limit and, thus, are not realistic for real cases. Nevertheless, I still use R1 and R2 because a q_r of 5.0 $\text{l}/(\text{s}\cdot\text{ha})$ refers to precipitation values that are not extreme, but respectably high. In interaction with other values, they might affect CSOs of more complex systems.

Finally, five different structures were added to the original system. A q_r of 2.0 is the limit of the standard design approach according to *ATV-A128* (1992). However, in practice sewers with larger q_r exist. Consequently, I add another set of three structures with storage in the modified artificial sewer model, where q_r is increased to 5.0 l/(s·ha). For structures without storage the maximum q_r used in *Drechsel* (1991) is 30 l/(s·ha). I included two additional structures without storage with increased q_r of 50 and 75 l/(s·ha). The reason is a better representation of extreme values, and thus the ability to compare the results of the indirect validation using the sewer and the results of IDF-curves.

6.3. Implementation of SWMM

In this section, the implementation of the individual modeling steps in Storm Water Management Model (SWMM) are explained in more detail. This is necessary because different implementation techniques can affect the validation results.

6.3.1. Runoff generation

In SWMM, the parameters influencing the runoff generation are depression storage, evaporation, impervious area and infiltration parameters. All parameters affect the runoff by reducing the effective precipitation, but in slightly different manners. The depression storage cuts precipitation values at the beginning of a spell, whereas infiltration may reduce all values, but is larger in the beginning of a spell than at the end. If the depression storage or infiltration are high, the small values are cut, and cannot contribute to runoff anymore. However, for an indirect validation the effect of the entire precipitation distribution should be validated, and thus losses due to infiltration and depression storage of the artificial model should be realistic, but rather small. In this study, the infiltration losses are set to zero by applying an impervious area of 100 %. I choose the depression storage to be 0.7 mm, which is at the lower bound of a realistic bandwidth for impervious areas. The evaporation parameters are of minor importance, but for a better comparability with the KOSIM model, I apply the same approach. That is, the evaporation is set to 600 mm/year including a daily and annual cycle as explained in *itwh* (2009).

6.3.2. Overland flow

The overland flow in SWMM can be modeled in different ways. In order to get the most appropriate one for the indirect validation I undertake the following preliminary considerations.

If two events with the same precipitation volume are regarded, the following holds true. The more non-uniform the precipitation intensity distribution is, the more likely CSOs occur as q_r is exceeded. In other words, the more uniform the flow, the least critical it is with respect to CSOs, and thus the less important the precipitation structure of the driving process

is. However, with an indirect validation one tries to infer the characteristics of the precipitation structure. Therefore, the modeling of overland flow should be chosen such that the hydrograph is attenuated the least.

On the other hand, if zero, or almost zero, attenuation is considered, the indirect validation of the artificial sewer may show existing deficiencies in the precipitation structure, but they can be entirely irrelevant for real sewer systems. This is because the hydrograph is always smoother compared to the hyetograph due to natural storage and retardation processes. In this respect, the indirect validation should help to find the most important characteristics of the precipitation time series by smoothing the hydrograph, and so filtering unimportant properties. Consequently, the overland flow should be modeled such that the hydrograph is attenuated least, but, additionally, is still in a realistic range.

One possible implementation of the overland flow with less attenuation is by taking the runoff coefficient method with a high runoff coefficient. However, this approach has some disadvantages. Despite of being very simple, it has been developed for peak flow estimations of sewer design and not for modeling CSOs. Furthermore, the attenuation is a simple trimming of the time series by a given fraction without a representation of translation and attenuation, and thus without any smoothing.

Another method, which I applied in my study, is based on the continuity and Manning equation. The advantage is that the time series is smoothed due to translation and attenuation effects, and so unimportant characteristics of the time series can be filtered. The smoothing effect depends on the parameters which are the Manning coefficient, the slope and the width of the catchment. All three parameters influence the hydrograph in a similar manner. Therefore, according to *Rossmann and Huber (2016)* the Manning coefficient as well as the slope should be fixed and the calibration should be done by varying the width. An increase in the width for the same catchment area decreases the length of the flow path and, thus, reduces the attenuation of the peak flows. Consequently, a large width reflecting a short flow path smooths the hydrograph the least and shows the underlying precipitation characteristic the best.

The question rises how the parameters should be chosen for an artificial sewer system. Usually, the calibration is performed using a measured hydrograph. Obviously, this method cannot be applied to a synthetic system. Therefore, I fix the Manning coefficient and the slope to realistic values emphasizing fast runoff, while I alter the width. As this is a crucial aspect, I will investigate the selection of the width in a more detailed sensitivity analysis later in Chap. 6.4.

6.3.3. Sewer flow routing

In this study, I apply the steady state approximation for sewer flow routing instead of dynamic or kinematic routing for three major reasons. First, hydrodynamic routing has to be used if not only flow rates, but also depths, are of interest. This is the case for hydraulic dimensioning of sewer networks and surface flooding, but not for long term sewer system simulations.

Second, dynamic routing is also necessary if effects like backwater, pressurized flow or flow reversal are important. These are more relevant for smaller slopes, larger in-pipe storage volumes as well as for smaller throttle discharges causing backwater (*Schmit and Hahn, 1986*). However, according to them it depends highly on the regarded sewer system if some simplifications can be done without significant deviations. Therefore, the actual impact is related to the specific layout including pipe diameters, the number and setting of manholes or the location and design of CSO structures. Therefore, it is very difficult to define a representative artificial sewer system including these effects.

Finally, in a preceding case study of *Brockerhoff (2015)* the influence of both routing methods was analyzed. The applied sewer system was rather flat, and thus, differences in the routing methods should occur. Sewer simulations with measured precipitation as well as with synthetic data of *NiedSim* were performed. *Brockerhoff (2015)* analyzed the deviation of measured and synthetic precipitation with respect to different CSO statistics for three overflow structures. A comparison between the routing techniques showed differences which depend on the regarded overflow structure and CSO statistic. Interestingly, the deviations in the simplified steady state routing were systematically larger than in the complex hydrodynamic routing, and thus errors in the synthetic time series of *NiedSim* are more important for simple routing. She concluded that steady state is the better routing technique for an indirect validation of precipitation data.

6.3.4. Post processing

For combined sewer system simulations, the mean annual CSO volume is the most important parameter because it is used in *ATV-A128 (1992)* as design parameter. For an indirect validation one should also check if the mean annual CSO volume is correct for the right reasons. One could either investigate the entire distribution of all CSOs contributing to the annual sum or one analyzes additional statistics like the frequency and duration of CSO events. Here, I use the latter approach because it is more robust and one tends less to over-interpretations of a too detailed analysis.

In contrast to *KOSIM*, *SWMM* does not automatically provide CSO statistics as frequency [events/year], duration [hours/year] and volume [$\text{m}^3/(\text{hectare} \cdot \text{year})$]. Consequently, they are conducted in a post process from the *SWMM* output of the spill flow of each structure. As the resolution of the input precipitation time series is 5 min, I choose the reporting time step of *SWMM* to be 5 min as well. The CSO duration is simply calculated by counting the time steps where spill overflow occurs. For the frequency, however, events must be defined which is crucial as already explained for precipitation in Chap. 2.2.2. In my study I defined the events directly related to the reported start and end of an overflow time period, without considering the emptying time of each CSO structure as applied in *KOSIM*. This way, the internal intensity structure is reflected better, but on the cost of separating events that are dependent on each other. This issue affects only the mean annual frequency, which is one reason, among others explained later, why I will barely use the frequency in my study. Finally, I set the event interim time to 15 min instead of 5 min in order to avoid numerical issues. The CSO statistics shown herein are all mean annual values.

6.4. Sensitivity analysis of the artificial sewer

The definition of the catchment width is the most influencing parameter with respect to CSOs of the artificial sewer apart from V_s and q_r . In order to find values which are appropriate for an indirect validation, I performed a sensitivity analysis. For this purpose, the Manning coefficient is chosen to be 0.013, which represents the roughness of paved surface. The surface slope is set to 2.5 % reflecting a medium gradient. The range of the width is varied from 30000 – 200 m representing a length of overland flow l_o of 33 – 5000 m for an area of 100 ha. According to *Rossman and Huber* (2016), the maximum length of l_o is between 33 – 150 m. Therefore, the maximum length of 5000 m in the sensitivity analysis is very high, and not realistic for overland flow of real catchments. However, I included higher values to account for additional attenuation in the pipes of the sewer. As the sewer flow routing is steady state, the shape of the hydrograph is entirely defined by the overland flow transformation process. A scenario, where the attenuation in the pipes is neglected, would reflect a CSO structure located directly after the inlet of the sewer, which is usually not the case. The larger widths should account for CSO structures located farther away from the inlet of the sewer system, where the hydrographs are more attenuated.

For the sensitivity analysis, the sewer systems are simulated using measured precipitation data of Freiburg from 1961 – 1980 (Period I). Fig. 6.2 shows the absolute CSO statistics for different overland flow lengths. Before the influence of the width is pointed out, some general aspects of the artificial sewer are explained.

From the figure, one can notice that the overflow frequency, duration and volume are each highly negatively correlated with an increased V_s or an increased q_r . This behavior shows the filtering characteristic of the different CSO structures explained in Chap. 6.1. Furthermore, it should be pointed out that the absolute values of structures with large V_s or q_r are quite low. This is an important aspect for the evaluation of the relative deviations later in my study. The smaller the absolute value, the higher the relative deviation can be.

Coming back to the sensitivity analysis of the overland flow path, one can clearly notice the high influence of the parameter on the CSO characteristics in Fig. 6.2 (left panel). The larger the length, the more the flow is attenuated which leads to lower CSO frequencies and volumes because the flow can pass the throttle more often without causing CSO. For the overflow duration, the characteristic depends on the structure. For structures without storage volume (R1 – R6), the duration behaves like frequency and volume. However, for structures with storage, in particular if q_r is small (11, 31, 51), the duration increases with increasing length. The reason is that these structures fill up quickly and the attenuated, but at the same time prolonged precipitation spells increase the CSO duration.

The absolute values, although highly influenced by the flow path, are of minor importance for the indirect validation in my study. For the indirect validation, the deviations between two precipitation time series with respect to their CSO behavior is analyzed to infer on the quality of the synthetic time series. Consequently, the question rises if a comparison of two time series depends on the length of the overland flow l_o at all.

In order to answer this question, I used two time series: a series with a measured intensity structure and the same one with block rainfall, where the intra-hourly intensity structure is

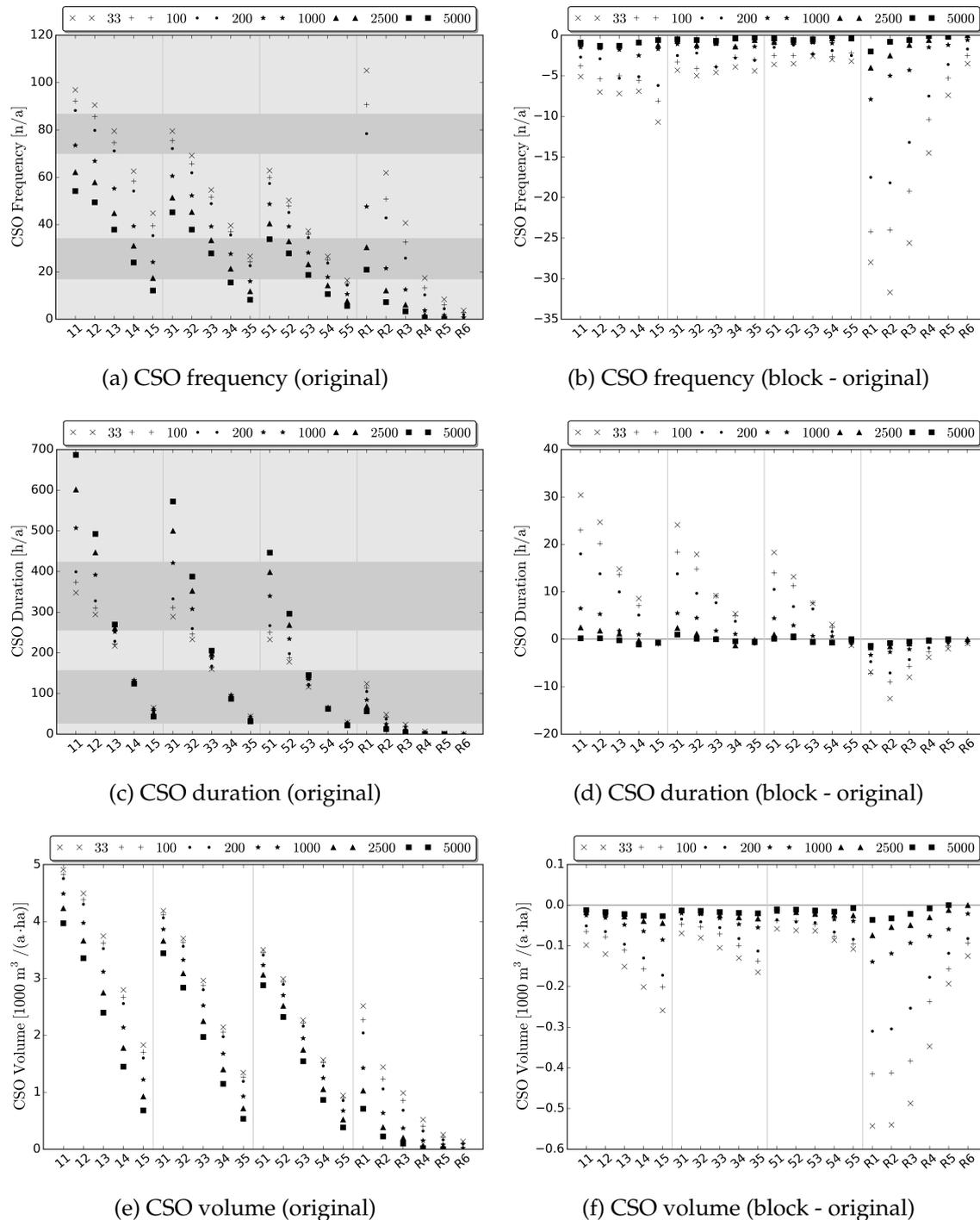


Figure 6.2.: Left: Results of the CSOs of the original measured time series of Freiburg from 1961 – 1980. The different markers represent the overland flow lengths l_o . The light and dark gray shaded areas show the bandwidths of the five different classes for stormwater tanks according to *Brunner (1997)* (the overflow volume misses as it was not part of his study). Right: Deviation between original measured time series and same time series as block rainfall.

destroyed by setting all twelve 5 min values to the average 5 min intensity of the respective hour. A sensitivity analysis of l_o using these time series is appropriate as explained in the following. The higher the temporal resolution, the more the attenuation and smoothing effect of different l_o influence the hydrograph and the more important l_o is. Therefore, the largest effect can be obtained at the 5 min aggregation. On the other hand, the strongest simplification of a time series at the 5 min aggregation is the assumption of block rainfall. Hence, the influence of l_o on the indirect validation can be analyzed by comparing the CSO statistics of two time series, one with the original intensity structure and the other with block rainfall.

Figure 6.2 (right panel) illustrates the relative deviation between the original, measured time series and the same one with block rainfall. Apparently, the deviation of both time series with respect to CSO statistics can be considerable high and depends on l_o . The largest deviations occur for the shortest, but still realistic l_o (33 m). The longer the path, the less sensitive the CSO structures respond, and thus the less important the underlying precipitation structure is. Consequently, small l_o are necessary to investigate high resolution intensity structures.

Characteristics of rainfall that do not influence the CSOs of a sewer system with $l_o = 33$ m are supposed to be unimportant in practice. However, not all deviations occurring in an artificial system with $l_o = 33$ m do necessarily affect real sewers because the hydrograph can be more attenuated in there as shown in the next section.

6.5. Validation of the artificial sewer

I set up the artificial sewer according to the generally recognized regulations of technology of Germany for combined sewers according to *ATV-A128* (1992). Nevertheless, the synthetic sewer is based on several assumptions and simplifications, and thus the question rises: How well can the artificial test sewer reproduce the CSO statistics of the overflow structures of real sewers? For this purpose, I pursue two different approaches, a literature study and a real case study.

I compare the artificial sewer with existing sewer structures in Germany described in a study of *Brunner* (1997). He analyzed combined sewer systems in Baden-Württemberg by introducing a ranking method for bypass tanks and pass-through tanks. He ranked all structures according to overflow frequency and duration. Afterwards, he defined five classes of the ranked data with equal frequencies. The gray bandwidths in Fig. 6.2 (left panel) show the five classes of the pass-through tanks. As the data of the study shows only sewer systems in Baden-Württemberg, it might not cover the entire range of existing CSO structures with storage. However, it gives an idea of the large range of CSO statistics of existing structures, and it shows that this range is well covered by the artificial sewer.

For the comparison with a case study I use the sewer of Braunschweig in Germany (*van der Heijden et al.*, 2017). It is a flat sewer with 96 CSO structures without storage, in addition to one CSO structure with a pass-through tank ($V_s = 21.2 \text{ m}^3/\text{ha}$). Most of the CSO structures

without storage are simple weirs that do not contain particular throttles. The CSOs are mainly caused by backwater, and therefore the actual runoff discharge rates q_r cannot be explicitly defined. Consequently, I cannot compare structures in the case study to the ones of the artificial system using V_s and q_r . Instead the structures are classified by a comparison of normed event-based CSO statistics. The analysis of event-based CSO statistics is a major difference to the previous and following examinations, where I investigate mean annual statistics.

For this purpose, the CSO statistics of each event and of each structure of both sewers are standardized by the sum of the respective CSO statistic, that is

$$x_{k,i} = \frac{X_{k,i}}{\sum_{i=1}^n X_{k,i}} \quad (6.1)$$

where $X_{k,i}$ is the CSO statistic of a structure k for an event i , $x_{k,i}$ is the respective normed statistic and n is the total number of events. $x_{k,i}$ can be regarded as an individual response time series of a CSO structure caused by the precipitation events. The response depends on the location of the structure in the sewer and on the specifications of q_r and V_s . The statistics used in my study are frequency, duration and volume of the CSO event. As the response time series of the structures of the case study and the synthetic sewer are standardized, they can be compared by using a quality measure to quantify their differences, for example, the Nash-Sutcliffe efficiency coefficient (NS) (Krause *et al.*, 2005). I use this quality measure to find the most corresponding artificial structure for each structure of Braunschweig.

Some questions can be answered by this matching procedure: First, are some CSO statistics better represented than others? Second, which simplified artificial structures actually reflect structures in the case study, and how good is the quality of the representation? Finally, is the high sensitivity of the synthetic sewer ($l_0 = 33$) necessary for real cases?

For this purpose, I separate a precipitation time series of 22 years (1985 – 2006) of Braunschweig into individual rainfall events using a 5 min intensity threshold of 0.1 mm, a minimum event depth of 1.5 mm and an interim time of 8 h assuring independence of the events for this sewer. The resulting $n = 487$ events are used to run the sewer model of Braunschweig as well as the synthetic sewer with $l_0 = 33$ m. Only 25 structures of Braunschweig showed considerable CSO activities. I defined the CSO activity as structures that contribute at least 0.5 % to the total overflow volume of the sewer system. I compared each event series of the 25 structures of Braunschweig with the event series of all 21 structures of the synthetic sewer. The NSs for all 525 combinations are calculated according to

$$NS = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (6.2)$$

where y_i is the normed CSO statistic x_i^k of the artificial sewer, \bar{y} reflects the mean of x_i^k of the synthetic sewer and \hat{y} is the x_i^k of the real sewer. A NS = 1 shows a perfect fit, that

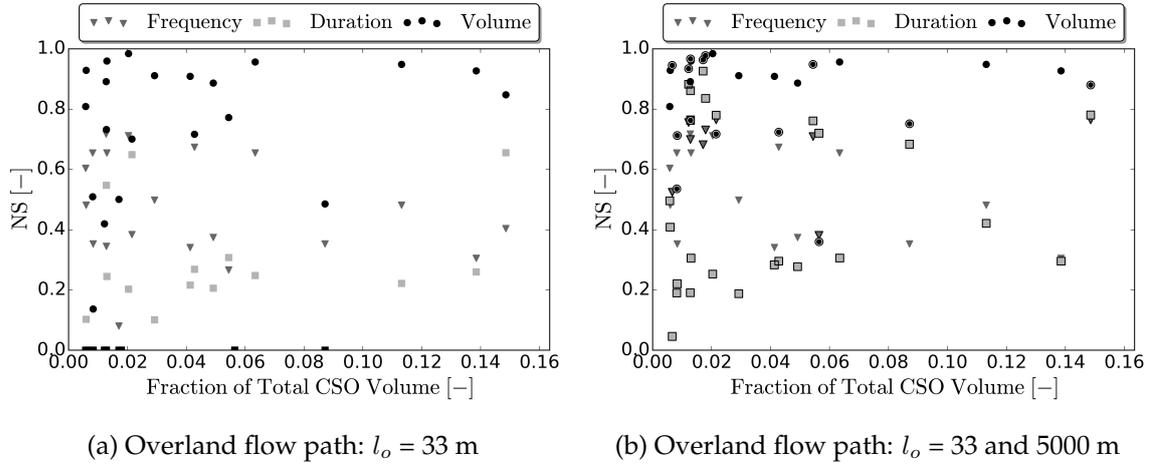


Figure 6.3.: NS of the best matching CSO structures of the artificial sewer system for different CSO statistics. The NS is plotted against the fraction of the CSO volume of the respective structure from the total volume of all structures. The framed markers reflect $l_o = 5000$ m. Remark: The y-axis is limited to zero and all negative NSs are plotted at $y = 0$.

is, the CSO event series of artificial and real sewer are identical. In other words, the event series of an overflow structure in the real case can be perfectly represented by the structure of the synthetic sewer. A $NS \leq 0$ reflects the case if the event series of the artificial case is not better (or even worse) than just using the mean CSO statistic of the synthetic case for the representation of the real sewer.

In the first analysis, I analyzed the maximum possible quality of the fitting for each of the 25 structures of Braunschweig with the artificial sewer. That is, the highest NS of each structure of Braunschweig is regarded, independently which synthetic structure provided the best matching. The analysis is individually done for all three CSO statistics. The results are shown in Fig. 6.3a. The NS is plotted against the fraction of total CSO volume of the respective structure to investigate if the quality of fitting depends on the amount of CSO. The different markers are the NSs of the CSO frequency, duration and volume of the structures in Braunschweig. For example, the structure with the highest overflow volume (contributing approximately 15 % to the total volume of all structures) shows an NS of 0.84 (in volume), 0.65 (in duration) and 0.41 (in frequency). The figure illustrates that the CSO volume of the case study can be represented much better compared to the frequency or the duration. Furthermore, I demonstrate that the representation of the CSO statistics of a structure is independent of its fraction of the total CSO volume. Large CSO structures are neither better nor worse represented than small sewers.

In the next investigation, I match and classify the 25 structures of Braunschweig to the artificial structures showing the highest NSs (Tab. 6.2). The first line describes the number of structures of the case study that are represented by the artificial structures. The second line shows the average quality of the matching as a mean NS. For example, with respect to the CSO volume, structure 35 of the artificial sewer represents 4 of the 25 structures of

CSO statistic	Structure of artificial sewer							
	15	34	35	55	R3	R4	R5	R6
Frequency								
Number	1	0	5	8	2	1	1	7
Mean NS	0.30	-	0.55	0.30	0.59	0.72	0.6	0.15
Duration								
Number	0	0	1	22	0	2	0	0
Mean NS	-	-	0.26	-2.15	-	-1.46	-	-
Volume								
Number	1	3	4	14	0	0	1	2
Mean NS	0.91	0.91	0.95	-0.2	-	-	0.81	0.32

Table 6.2.: Classification of the CSO structures of Braunschweig to the best matching artificial sewer structures with $l_o = 33$ m. The first line of each statistic shows the number of matched structures in the case study. The second line reflects the quality of the matching indicated as mean NS.

Braunschweig the best, with a mean NS of 0.95 (calculated using the 4 individual NSs). It is noticeable that all synthetic structures occurring in the table have a high q_r . Although the q_r of the structures of Braunschweig are not explicitly known, the result should not be surprising. The majority has no storage, and thus the q_r must be high to avoid CSOs. Another interesting result is that the structures 55 and R6 show the highest number of matches, but at the same time have the lowest NS. This indicates that structure 55 and R6 do not represent structures of the real sewer well, but are only chosen among the 21 artificial structures because all others fit even worse.

The artificial sewer in the case of $l_o = 33$ m is a very fast responding system, where the flow is hardly attenuated and smoothed. Therefore, the individual structures are rather sensitive and CSOs occur frequently. The system of Braunschweig, however, is a flat system with relatively long flow times, and thus the flow peaks are less pronounced and CSOs occur less frequently. Consequently, the structures of Braunschweig, that are not well represented, are related to synthetic structures showing very low CSO activities with a high q_r (R6) or a high V_s and q_r (55). In other words, the artificial sewer is too sensitive to be fully representative for Braunschweig.

In order to demonstrate this hypothesis, I extend the existing sewer to 42 structures, where half of them are connected to an area with $l_o = 33$ m (as before) and the other half to an area with a path of 5000 m. The configurations of the individual structures are the same as before. For the extended artificial sewer system I apply the same matching procedure as explained above. Figure 6.3b shows a clear improvement of the representation. The classification of the best matchings (Tab. 6.3) illustrates that many structures of Braunschweig are not very sensitive and are better represented using $l_o = 5000$ m. However, some structures respond very fast, such that the artificial sewer with $l_o = 33$ m is still necessary for a good representation. A further increase of l_o to more than 5000 m changes the overall picture of the overflow characteristics only slightly. It is not surprising that the artificial sewer cannot per-

Statistic	Structure of artificial sewer									
	15	34	35	55	R3	R4	R5	R6		
Frequency										
Number	1	0	5	2	2	1	1	3		
Mean NS	0.30	-	0.55	0.48	0.59	0.72	0.60	0.45		
Duration										
Number	0	0	0	0	0	0	0	0		
Mean NS	-	-	-	-	-	-	-	-		
Volume										
Number	1	3	4	2	0	0	1	0		
Mean NS	0.91	0.91	0.95	0.94	-	-	0.81	-		
	14*	15*	34*	35*	54*	55*	R1*	R2*	R3*	R4*
Frequency										
Number	0	4	0	0	0	1	0	3	2	0
Mean NS	-	0.75	-	-	-	0.70	-	0.72	0.45	-
Duration										
Number	0	1	0	8	2	0	1	3	5	5
Mean NS	-	0.3	-	0.52	0.28	-	0.3	0.58	0.69	0.33
Volume										
Number	1	1	3	2	0	0	0	3	2	2
Mean NS	0.72	0.95	0.79	0.95	-	-	-	0.9	0.65	0.62

Table 6.3.: Classification of the CSO structures of Braunschweig to the best matching artificial sewer structures with $l_o = 33$ m and $l_o = 5000$ m (the latter is indicated with *). The first line of each statistic shows the number of matched structures in the case study. The second line reflects the quality of the matching indicated as mean NS.

fectly represent the statistics as it cannot reflect effects related to CSO structures connected in series.

Several conclusions can be drawn from this section. First, the artificial sewer can represent the CSO volume well, while the CSO frequency and in particular the CSO duration are more difficult to be reflected. Second, the artificial sewer with $l_o = 33$ m is very sensitive and, thus, shows the quality of the precipitation structure the best as explained in 6.4. However, real sewers exist that are less sensitive and, consequently, are better represented by $l_o = 5000$ m. The conclusion with respect to synthetic precipitation is, that erroneous time series might be detected by a sensitive artificial sewer, but are likely to be irrelevant for real sewers if they respond rather slow on precipitation events. This is the case for sewers, where the flow is highly attenuated and smoothed, for example, in flat sewers, in sewers with a long flow time, or in oversized sewers with a lot of additional storage capacity.

As the sensitive sewer with $l_o = 33$ m and the robust sewer with $l_o = 5000$ m represent different characteristics of real sewers, I use an artificial sewer with both configurations for the indirect validation in the rest of the study.

7. Preliminary investigations of the indirect validation using the artificial sewer

The artificial sewer is intended to be a tool for the validation of precipitation time series. However, as it is an artificial sewer, the absolute values of the simulations can hardly be verified or compared with measured sewer data. Consequently, I investigate the relative deviations in the CSO statistic x of the artificial sewer according to

$$x = \frac{X - X_{ref}}{X_{ref}} \quad (7.1)$$

where X is a CSO statistic (volume, duration or frequency) of the time series to be analyzed with respect to the CSO statistic of a reference data set X_{ref} . This way, the impact of differences in the rainfall input data on the application can be shown.

The investigation of relative deviations in the CSO statistics instead of absolute values has two advantages. As the focus in my study is on the temporal structure of the precipitation time series, the marginal distribution should not influence the validation much. However, different marginals of various locations or time periods can highly influence the CSO statistics. For example, a dry year will show smaller absolute CSO statistics compared to a wet year, even if the temporal structure is comparable. When relative values are used, the time series can be compared without the effect of different marginals. Furthermore, the deviations in the CSO frequency, duration and volume can be compared among each other although their absolute values have different units.

Nevertheless, relative values come along with a true short coming. Without considering the absolute values, they can easily be misinterpreted. In particular, if the absolute values are very small, the relative deviations can increase very fast and pretend a tremendous error, while their absolute differences are almost negligible. To judge the relative deviations, I use confidence intervals based on measured data. If the uncertainty of the measured data is expressed as relative bandwidths, the deviations of synthetic time series can be put into perspective by a comparison with the relative confidence intervals. Therefore, large relative differences between synthetic and measured time series are more tolerable, if the relative observed uncertainty is also high. The confidence intervals also define the limits of model accuracy for the simulation of synthetic precipitation because one cannot expect a model based on observations to be better than the input data.

In the first section (Chap. 7.1), I will analyze the sampling variability and obtain confidence intervals of the mean annual CSO output. In the second section (Chap. 7.2), I will examine the characteristics and behavior of the artificial sewer and try to infer the relevant

parts of the precipitation time series for the sewer simulation. In both sections, I will use only measured time series that are slightly manipulated depending on the respective objective and not simulated data.

7.1. Identification of sampling uncertainty

In Chap. 3, I analyzed the statistics of observed precipitation data, and showed and explained different sources of uncertainty including measurement errors. In this section, the influence of the uncertainty on the application, the artificial combined sewer, is described. I consider the influence of sampling uncertainty related to spatial and temporal variability as well as inhomogeneities due to different measurement devices. I do not investigate the measurement errors here due to the reasons I explained in Chap. 3.

This study focuses on the generation of point time series of single locations and not on the simultaneous generation of time series with a spatio-temporal correlation. Therefore, I assess the spatial variability by comparing different locations using the same time period to obtain systematic deviations in the sewer caused by local climate conditions. For the temporal variability, I compare the sewer output of different time periods at the same location. The latter task should be done with caution because long time series must be analyzed in order to get the entire spectrum of possible precipitation events. The shorter the time period, the more likely it is that the (short) samples are not representative for the underlying population. However, long time series are more vulnerable to inhomogeneities such as a change in the natural processes (e.g. climate change) or artifacts related to varying measurement techniques.

7.1.1. Assessment of temporal uncertainty

Precipitation can be seen as a stochastic process where each realization of the underlying population shows a high bandwidth of possible outcomes. However, the long term average of individual years approaches to the mean statistics of the population, and the longer the time series is, the more stable the average of the series is. As extensively shown in literature (e.g. *Xanthopoulos, 1990; Haberlandt, 1996*), several years are necessary in order to achieve robust estimates of the mean annual characteristics. According to *ATV-DVWK-M165 (2004)* at least 10 years of precipitation data must be used for applications of urban hydrology.

In order to estimate the natural variability with respect to the CSO statistics, I apply a bootstrap approach. The method enables to obtain an additional number of samples that have the same size as the original sample (bootstrap samples), where the single outcomes are drawn with replacement from the original sample. The set of bootstrap samples reflect the same characteristics as the original sample, but with an increased sample size, such that confidence intervals can be calculated. Even if a rather small subset (10 of 37 years) is taken, the total number of possible combinations with repetition sums up to $4.7 * 10^9$ possibilities. Nevertheless, as it is a resampling approach, it contains only information about the underlying observed data. Characteristics that were not measured, cannot be reflected

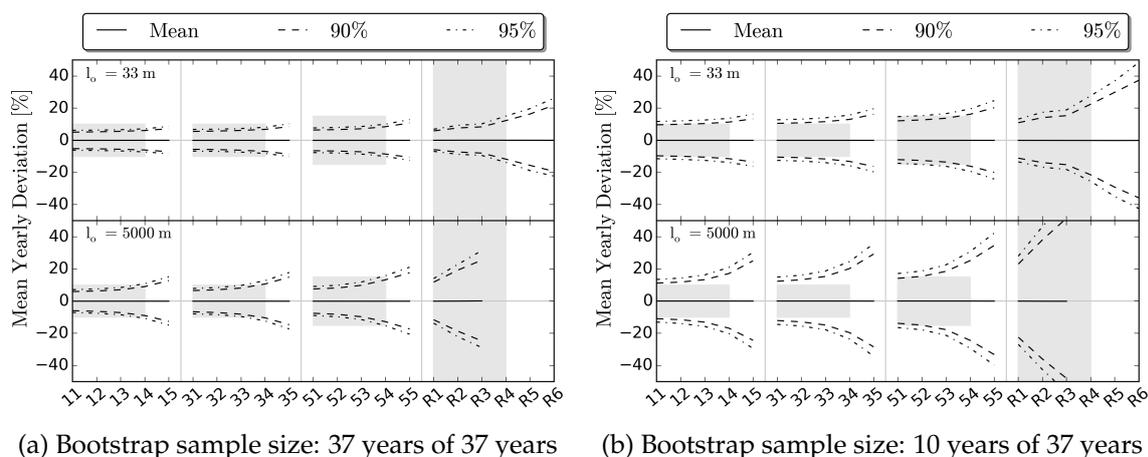


Figure 7.1.: The mean, the 90 % and 95 % confidence interval of the mean annual deviation of 10000 bootstrap samples from the original data of Freiburg for two different sample sizes of the CSO volume. The gray bandwidth shows the acceptable deviations according to *Drechsel* (1991).

with a bootstrap. For more information about bootstrap methods I refer the interested reader to *B. Efron* (1986).

An assumption of the bootstrap is stationarity, that is, all necessary information about the underlying population is included in the observed sample and does not change in time. Therefore, all future observations can be obtained by a resampling of observed outcomes from the past. In that perspective, the confidence intervals obtained from the following analysis are likely to be too narrow and, thus, are too optimistic as non-stationary effects like climate change are not considered.

I use the annual CSO statistics of Freiburg of all 37 years for the bootstrap. I draw $n = 10000$ bootstrap samples from each CSO statistic and calculate the relative deviation from the original statistics used as reference according to Eq. 7.1. The mean, the 90 % and 95 % confidence interval of the resulting relative deviation of 10000 bootstrap samples are shown in Fig. 7.1a. Figure 7.1b shows the bootstrap bandwidth of a reduced sample size using only 10 years of data which are the necessary minimum number of years according to *ATV-DVWK-M165* (2004). For this purpose, I apply a double bootstrap approach. At first, I draw $m = 100$ subsamples with replacement from the 37 years of the original data. Afterwards, I apply the bootstrap approach explained above (using $n = 1000$ bootstrap samples) to each subsample to observe the variability within 10 years. Finally, I calculate the relative deviations of the n bootstrap samples to the respective subsample. The mean and the confidence intervals shown in Fig. 7.1b are obtained using all 100×1000 bootstrap samples. In both figures, additionally, the gray bandwidths illustrate the tolerable deviations defined according to *Drechsel* (1991) (Tab. 7.1). Unfortunately, he did not describe how he determined his bandwidth of acceptable deviations.

As expected, the reduced sample size of 10 years shows an increased uncertainty compared to the 37 years. The two bandwidths can be interpreted as follows. The approach

Structure	B11 – B14	B31 – B34	B51 – B54	RR1 – RR4
Deviation	±10 %	±10 %	±15 %	±50 %

Table 7.1.: Acceptable bandwidth of deviations for different overflow structures according to *Drechsel* (1991).

using 37 years represents a long, but rarely available series that shows only small deviations in the confidence interval of around 5 – 10 %. The deviation of simulated time series needs not to be smaller than this interval because the measurements cannot provide more precise results. Therefore, the confidence interval of the 37 years of data can be seen as the actual target value for the synthetic precipitation. However, from a practical point of view, most measured time series are shorter than 37 years. Therefore, simulated data could be as uncertain as shown by the larger confidence intervals of the 10 year bootstrap sample size in order to meet still the required precision according to *ATV-DVWK-M165* (2004).

The bandwidths defined by *Drechsel* (1991) fit quite well to the confidence intervals of my study, except for structures without overflow (R1 – R6). Here, the bandwidth of him is much less restrictive compared to the bootstrap result. However, if one considers the ecological impact on the receiving water, the less confined bandwidth of *Drechsel* (1991) might be reasonable, because large relative uncertainties are less problematic for small absolute CSO values.

The uncertainty of the CSO duration is slightly larger (of the frequency smaller) compared to the volume. However, the deviations behave similar and, therefore, are not explicitly shown here. I use the confidence intervals of the CSO statistics of the full data set (e.g. 37 years for Freiburg) in this study for comparison, classification and as a benchmark for the hourly optimization of synthetic time series.

7.1.2. Origins of temporal uncertainty

In the last section I evaluated the effect of natural variability of precipitation time series. Now, the question rises what actually causes the indicated uncertainty. Does the marginal distribution of rainfall values contribute more to the uncertainty or their temporal order (structure)? This is interesting to know because in the simulation algorithm of *NiedSim* the marginal distribution of the 1 h precipitation values and the temporal structure are generated independently (Chap. 4.4).

In order to answer the question, I separate the marginals and the structures and recombine them year wise. The recombinations are done with a quantile-quantile-transformation (q-q-transformation) also known as quantile mapping. The advantage is that the manipulation of the time series is performed within its natural limits, which in particular means that the realistic temporal structure can be preserved. However, the resulting time series should not be seen as new, perfect samples of the natural precipitation process, as the assumption that the marginal distribution is independent of the structure is probably not true.

By applying the q-q-transformation the marginals of one year (target) can be replaced with the marginals of another year (source). For this purpose, the values of both time series

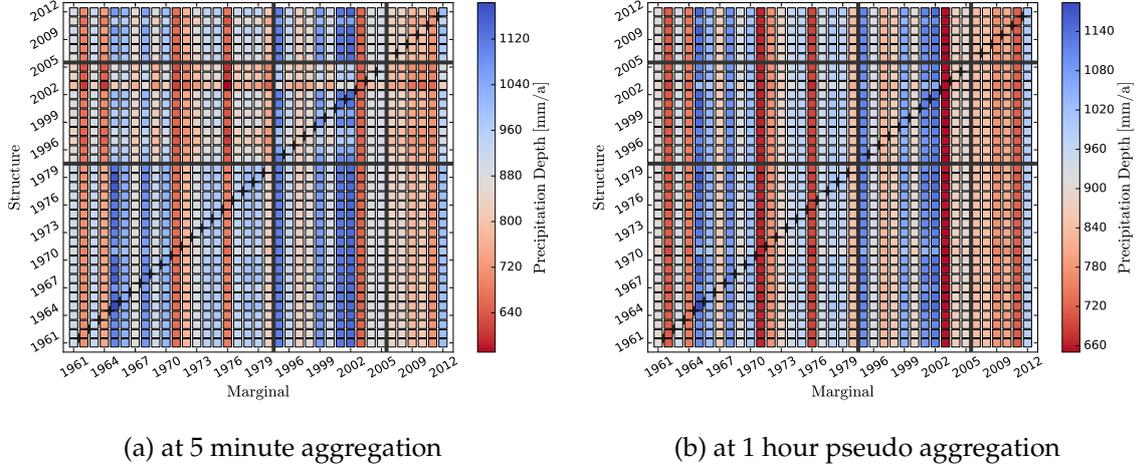


Figure 7.2.: Absolute values of the q-q-transformed precipitation depth of Freiburg for two different kind of q-q-transformation procedures.

are sorted by size and then paired by their quantiles. That is, I assign the largest value of the target data set to the largest value in the source; the second largest to the second largest and so on. Mathematically this can be expressed as:

$$x_{qq} = F_{tar}^{-1}(F_{src}(x_{src})) \quad (7.2)$$

where x_{qq} is the q-q-transformed data, F_{tar}^{-1} is the inverse of the CDF of the target data, F_{src} is the CDF of the source data and x_{src} are the values of source data.

As p_0 of both time series may differ, not all values larger than zero of one dataset have a respective value unequal to zero in the other dataset. If p_0 of the source data is larger than p_0 of the target data, the transformation can be applied straightforward because the target value is simply set to zero. In the other case, however, zero values must actually be transformed into precipitation values of the target time series to preserve the precipitation sum. For these values it is difficult to define a proper temporal location within the time series without changing its original structure because the location of zeros before the transformation are not unique. Therefore, the conversion of zero into non-zero values is not conducted. In order to keep the resulting volume error of the transformation as small as possible, the q-q-transformation starts by replacing the largest data pair. As soon as one value of the data pair is zero, the transformation stops and all remaining precipitation values are set to zero. The volume error introduced by this simplification is small as long as p_0 of both time series do not differ too much.

The result can be plotted in a $n \times n$ matrix with all combinations of structures and marginal distributions, where n is the number of years. Figure 7.2a illustrates the result of the q-q-transformation of the 5 min data of Freiburg. The matrix can be interpreted as follows: A comparison in the horizontal direction shows the variation of the years regarding the marginal distributions because on the x-axes the marginals change. Likewise, a vertical comparison indicates the variability of the structure. The gray underlined squares of the

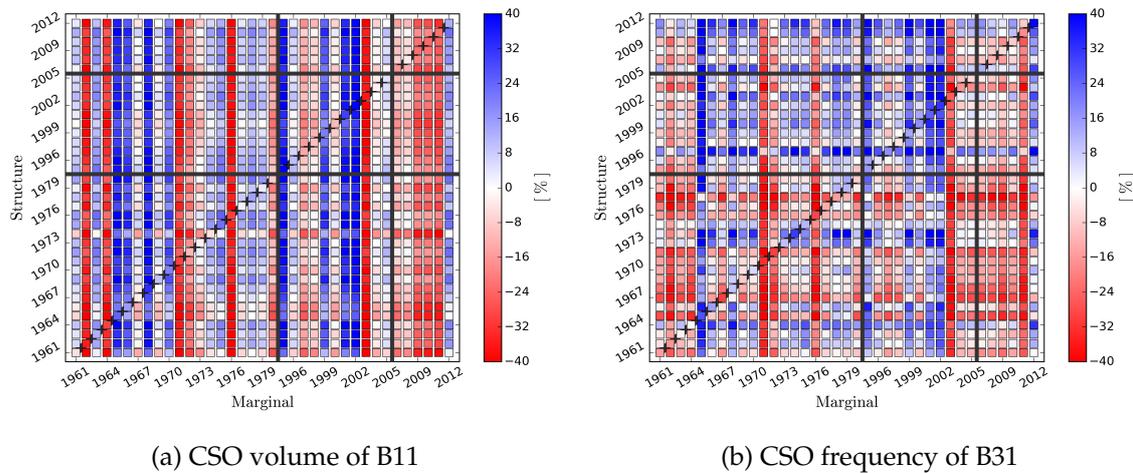


Figure 7.3.: Relative deviation of the q-q-transformed overflow statistics from the mean of the original overflow statistics (diagonal) of Freiburg.

diagonal are the original years and, thus, demonstrate the annual variability of measured data.

A problem occurs if the q-q-transformation is directly applied on the 5 min data. Figure 7.2a should show vertical lines of identical values, as differences in the structure should not influence the q-q-transformation of the annual precipitation depth. However, large deviations can be observed for many years due to large differences in p_0 . For example, very large deviations occur if the marginal distribution of the wet year 1965 is transformed to the time period of 1995 – 2005. As $p_0^{(5min)}$ of the time period I is systematically lower compared to period II and III (Tab. 3.2), a lot of precipitation values are cut in the q-q-transformation. This issue leads to a systematic underestimation of the annual precipitation depth and is intensified by the inhomogeneities of different measurement devices.

A non-conservative transformation of the precipitation volume is useless for the desired comparison of the influence of marginals and structure. Therefore, I apply a work around where in a first step the aggregated 1 h time series are q-q-transformed. Afterwards, the 5 min values of each hour of the original time series are rescaled with a factor defined by the relation of original and q-q-transformed hourly rainfall sum. The result is a q-q-transformed time series with the same, but scaled 5 min structure of the original data. Figure 7.2b indicates that the volume is well preserved except for very dry years that are transformed into very wet years.

I perform a sewer simulation using all 1369 q-q-transformed annual time series (37×37 years). In order to be able to compare different output parameters, I refer all absolute results to the mean of the original time series (the mean of the diagonal). Figure 7.3a is an example for a large influence of the hourly marginal distribution and a weak impact of the hourly temporal structure on the overflow statistics because one can observe pronounced vertical lines. Figure 7.3b is an example where both influences are similar. Both figures also show the high annual variation of CSO statistics, which underlines the necessity of taking long term means to get stable outputs.

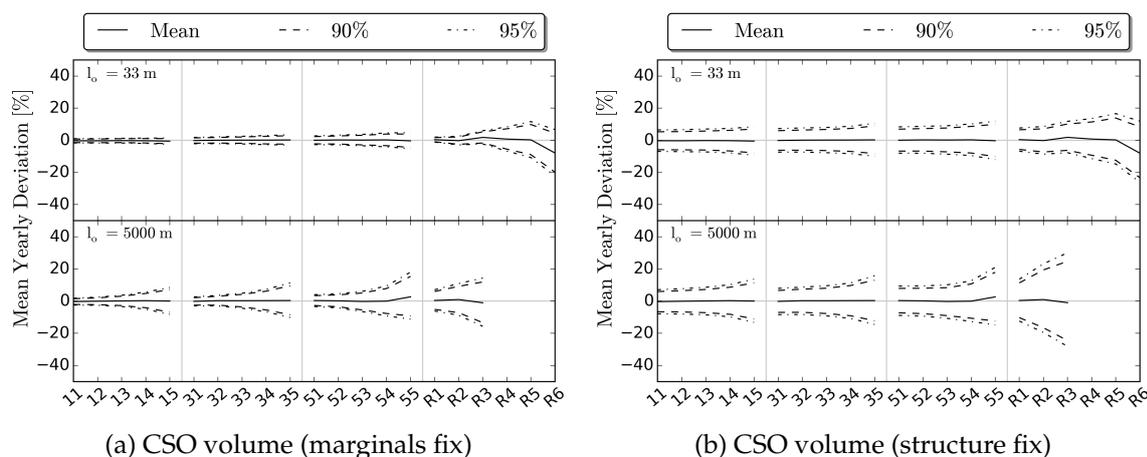


Figure 7.4.: Mean, 90 and 95 % confidence interval of the mean annual deviation of 10000 bootstrap samples from the original data of Freiburg for two different bootstrap samples. On the left, the marginals of each year are fixed and the structure is randomly chosen, on the right, it is the other way round.

Finally, I apply a similar bootstrap approach as explained before to show the part of the hourly precipitation time series (structure or marginal) that contributes more to the annual variability of the CSOs. The influence of both parts are examined individually by using the original marginals while varying the structures randomly, and the other way round. For example, the influence of the structure can be illustrated if each bootstrap sample contains the CSO data of all years with the original marginals $\{1961, 1962, \dots, 2012\}$, but with a random structure, for example $\{2007, 1963, 2007, \dots\}$. In the matrix one can think of choosing for each marginal in x direction a random structure in y direction. I refer the mean annual CSO statistic of each bootstrap sample on the mean CSO statistics of the original time series (the diagonal of the matrix). I draw 10000 bootstrap samples and calculate the mean, the 90 % and 95 % confidence interval of the relative deviations. Figure 7.4a shows the results for fixed marginals and random structures and Fig. 7.4b for fixed structures and random marginals.

The results of Fig. 7.4 indicate that the uncertainty of different hourly marginals have a higher impact on the overflow statistics compared to variability of the hourly structure. In other words, the structures of different years are rather similar. In particular, the CSO volume is almost independent of the annual structure. As the variability in the structure of different years has only a small influence on the CSO statistics, it is sufficient to reflect the structure on average rather than reflecting its inter-annual variability.

Not only the inter-annual variability of the marginals is larger, but also the impact on the CSO statistics. As shown before in Fig. 6.1b, the systematic reduction of the annual overflow volume compared to the other two time periods can be related to the rather dry period of 2007 – 2012 (Tab. 3.2). However, I would like to highlight that one cannot conclude that the structure is unimportant. The marginals are only more relevant if the structure is reflected realistically. Otherwise, large deviations in the CSO statistics can be obtained, even if the marginals are represented perfectly as I will show later.

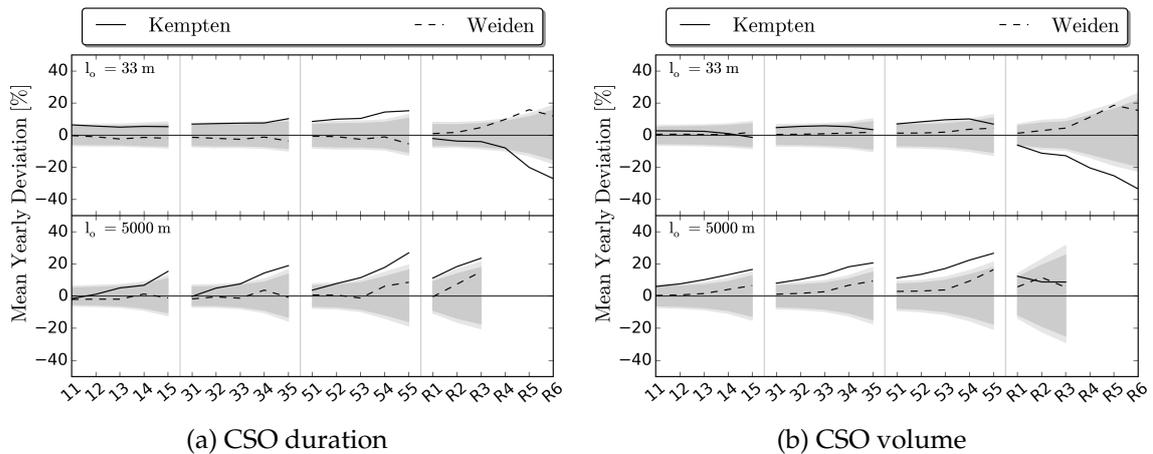


Figure 7.5.: Mean annual deviations of the CSO statistics of the mixed time series and original time series of Freiburg. The mixed time series consists of the marginals of Freiburg and structures of Weiden or Kempten. The gray bandwidths are the 90 / 95 % confidence intervals of 37 years of Freiburg obtained in Chap. 7.1.

7.1.3. Assessment of spatial uncertainty

In the last section I illustrated that the influence of the annual variation of the structure is lower compared to the variations of the marginals. If the marginal distribution dominates the structure, one could conclude that any kind of natural structure of a measured time series is valid to represent the local precipitation characteristics. In this section, I investigate this assumption by mixing the marginal distribution of one station with the structure of another station. I apply the mixing with a q-q-transformation as explained before.

I transform the marginals of the 37 years of Freiburg year wise on the 37 years of the same time period of Kempten. As for Weiden only 34 years exist, I use the respective 34 years of Freiburg. Each year consists of the marginals of Freiburg and the temporal structure of Kempten or Weiden of the same year. I perform a sewer simulation using each year and compare the result to the 37 years of original data of Freiburg by calculating the relative deviations.

Figure 7.5 illustrates systematic biases that can occur if the temporal structure of another station is used. For Kempten, where the climatic conditions are very different compared to Freiburg, the deviations of the CSO volume, duration and frequency (not shown) are larger than the 90 % confidence interval of the natural variability of Freiburg. Consequently, the temporal structure of different locations can have a significant influence on the CSO statistics. However, the spatial bias is still smaller compared to the temporal uncertainty of short time series of only 10 years (compare Fig. 7.1b).

In conclusion, although the temporal structure influences the CSO statistics to a less extent compared to the marginals or to short time periods, the structure is not interchangeable between different locations. Each location seems to have a specific structure based on its climatic condition. As for Weiden the influences are less pronounced, locations of similar

climate conditions might be found for which the temporal structure is interchangeable.

7.2. Important characteristics of precipitation with respect to combined sewers

In this section, I analyze the importance of the intra-hourly temporal structure on the CSO statistics as well as the influence of extreme values and the relevance of small intensities. All three sub-sections help to focus on the important parts of the generation and disaggregation of synthetic time series.

7.2.1. Influence of the intra-hourly intensity structure

The investigation of the intra-hourly intensity on the CSO statistics can show the general importance of the temporal structure with respect to different CSO structures. For this purpose, I analyze two differently manipulated measured 5 min time series. At first, I investigate a time series with block rainfall because this is the initial series for the NiedSim disaggregation. I obtain the block rainfall by redistributing the 5 min intensities uniformly within each hour. In the second time series, I randomly change the block rainfall using the NiedSim disaggregation algorithm (adding/subtracting small increments of precipitation), but without any optimization. The resulting time series exhibit a random intensity structure. Note that in both cases I change only the intra-hourly time series whereas higher time scales remain identical.

With both manipulated time series I perform a sewer simulation. I compare the resulting mean annual CSO statistics of each structure with the results of the original time series and plot them as relative deviations (Fig. 7.6 left panel). An extensive analysis of the influence of intra-hourly temporal structure follows, as I can illustrate some general characteristics of the behavior of the sewer system here. For a comparison with natural uncertainty I plot additionally the 90 and 95 % confidence interval of the 10 year time period for each CSO statistic as obtained in Chap. 7.1.

Block rainfall The transformation into block rainfall smooths the intensity distribution within one hour, and thus reduces intensity peaks. As explained in Chap. 6.4, the more attenuated and continuous the precipitation, the better CSO structures can handle the flow without causing an overflow. Consequently, the frequency and volume are reduced leading to an underestimation of the statistics, whereas the duration shows an overestimation.

The main influence of block rainfall can be related to structures without storage, where all CSO statistics are highly reduced and where the deviations can be larger than 50 %. However, one should keep in mind that the absolute values of the structures R3 – R6 are rather small. For structures with storage the block rainfall has relatively little influence on CSO statistics and the biases are in the 90 % confidence intervals of the natural uncertainty.

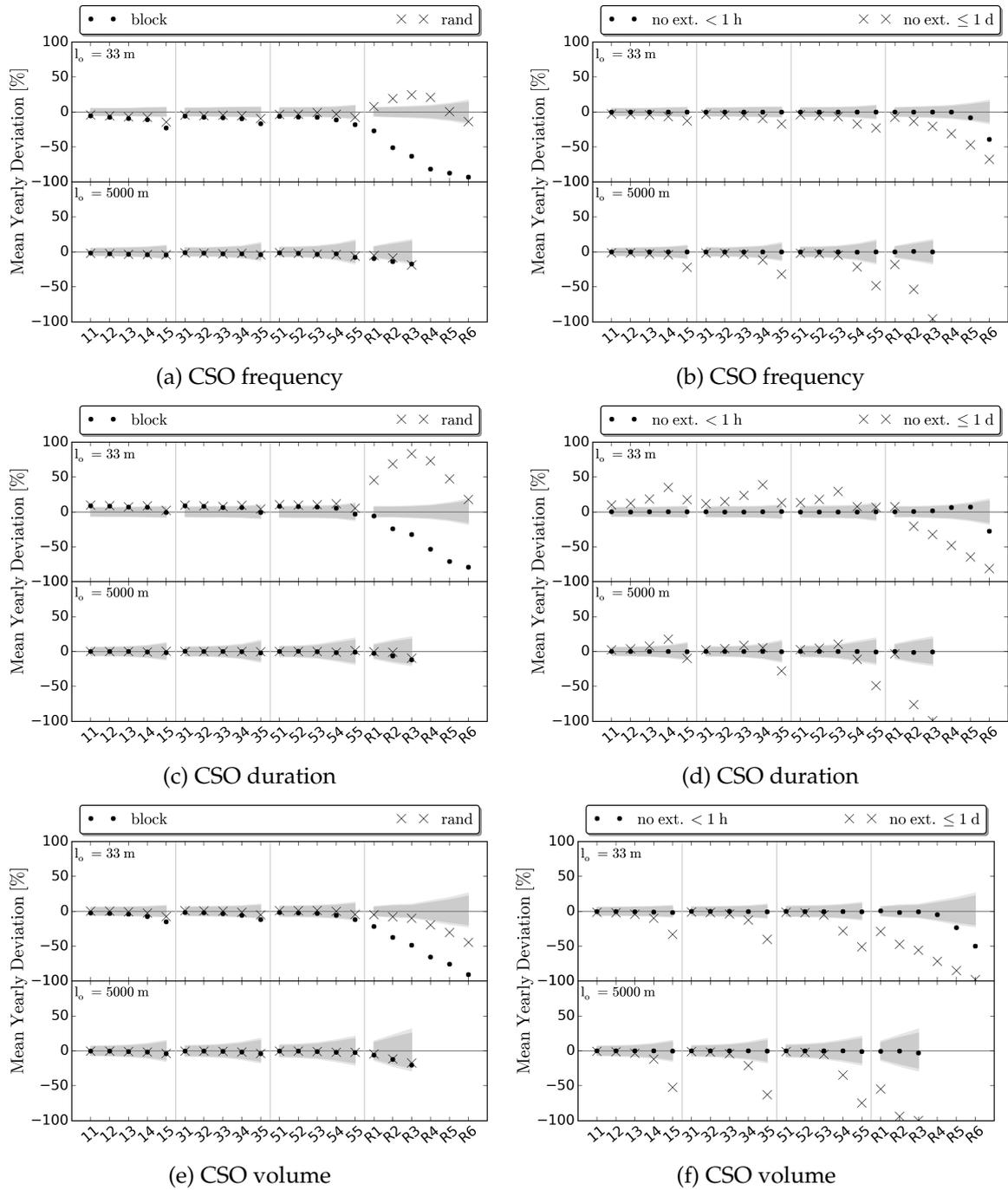


Figure 7.6.: Mean annual relative deviation of modified time series from measured originals. The gray bandwidths are the 90 / 95 % confidence intervals of 37 years according to Chap. 7.1. All results correspond to Freiburg (1961 – 1980). Left: Intra-hourly intensities are modified: Uniformly distributed (block) and randomly changed (rand). Right: Extreme values defined by the partial series ($n = 3$) are uniformly distributed. < 1 h: only intra-hourly scales; ≤ 1 d: all scales up to daily resolution.

The characteristics explained above are similar for different time periods and can also be detected for Kempten and Weiden (figures not shown). Therefore, the results can be assumed to be a general interaction of precipitation characteristics with different overflow structures independent of the regarded location or temporal uncertainty.

Random intensity structure The results of a time series with a random structure (Fig. 7.6 left panel) illustrates, that some of the deviations can already be reduced by a simple increase of the variation of the intra-hourly values. As already explained, the uniform distribution is least critical for the combined sewer system. Consequently, any disturbance of the block rainfall will increase the CSOs. However, the CSO duration is even more overestimated compared to block rainfall. This reveals that the larger intensities are scattered within the hour and do not form proper precipitation events with a pronounced start and end. Therefore, the challenge of the disaggregation is to force the time series to rebuilt intense values clustering together and forming events with a pronounced start and end.

In the literature, a correct temporal structure is usually compared to block rainfall, and consequently huge deviations are reported. For example *Zhu and Schilling* (1996) and *Drechsel* (1991) illustrate differences of at least 30 % in CSO volume using hourly block rainfall if structures without storage are considered. However, as shown above, such deviations can be reduced easily if not block rainfall, but randomly distributed intensities are applied. Therefore, the deviation of the random intensity structure (and not of block rainfall) from the original time series should be considered as benchmark, which I will do for the evaluation of the disaggregation later in my study.

In general, I can conclude that the influence of the intra-hourly intensity structure on combined sewer systems is less important. Only if short overland flow paths are considered, significant deviations can be observed compared to the natural uncertainty. The CSO duration is most sensitive to a correct intensity structure, whereas for the frequency and volume only CSO structures without overflow are critical. Nevertheless, an adequate disaggregation should be performed because otherwise small, but systematic biases can occur in CSO statistics.

7.2.2. Influence of extreme values

In the previous chapters, I assume that not only large, but also medium precipitation spells are relevant for combined sewer systems. Here, I investigate this assumption and the question, how much extreme values actually influence the CSO statistics. As extremes are usually analyzed by IDF-curves, the results will also show if both validation methods are redundant or independent.

For this purpose, I generate another two manipulated time series, where the extreme values of a measured time series are uniformly distributed. At first, I obtain all values which contribute to the partial series (≈ 3 events per year, compare Chap. 2.2.3) for different temporal aggregations. For each aggregation, I uniformly distribute the respective values at the 5 min aggregation. In this way, the intervention in the structure is kept minimal, but the

intensity of the IDF-curves is significantly reduced. I apply the procedure twice, for aggregations up to daily resolution {10, 15, 30, 45, 60, 120, 180, 240, 360, 720, 1440 min} and only for intra-hourly scales {10, 15, 30, 45 min}.

I perform a combined sewer simulation with both manipulated time series. Again, the relative difference of CSO statistics from the original time series are calculated. The results are illustrated in Fig. 7.6 (right panel). The influence of intra-hourly extreme values is only visible for the structures R5 and R6. It should be recapitulated that I implemented both structures in the original sewer systems of *Drechsel* (1991) to link the analysis of extreme values with the indirect validation by a sewer system. As the influence for all other structures is negligible, the intra-hourly extreme values can be changed without changing the CSO statistics. This is an important characteristic if one would like to adjust intra-hourly extreme values in a post-processing step as I will mention in the outlook.

If additionally extreme values of higher aggregations are considered, the influence is much higher, in particular for structures with large specific discharges ($q_r > 1 \text{ l}/(\text{s}\cdot\text{ha})$). The results show that extremes of larger scales comprise a substantial amount of precipitation volume, which contribute less to the overflow volume if the intensity is reduced by uniformly distributed values (Fig. 7.6f). On the other hand, the smoothed intensity in combination with structures including storage extend the CSO duration and lead to a strong overestimation (Fig. 7.6d). The smaller q_r , the more important the medium precipitation values are and the less the influence of the extremes is.

One can conclude, although extreme precipitation values cannot be ignored in CSO simulations, they do not dominate the results. In particular, for structures with storage, where the q_r is usually rather small, large and medium values are more important. Therefore, both validation techniques, that is, the IDF-curves and the indirect validation of the sewer system represent different parts of the time series.

Before I come to the final section that considers the influence of small values, I would like to share some thoughts about the CSO statistics used so far. The deviations of the time series with random intensity reveals that the duration is most sensitive to an erroneous temporal structure, whereas the volume and, in particular, the frequency are rather robust. This is not surprising as the relevant complexity of the rainfall events for CSO statistics reduces from duration to volume to frequency. For a correct CSO frequency, only a certain number of events need to occur, whereas for the volume, additionally, the cumulative sum must be accurate. The CSO duration, however, can only be reflected correctly if the rainfall events are long enough and at the same time large enough to cause overflow at all. In literature, usually only the volume is used for an indirect validation (*Andrés-Doménech et al.*, 2010; *Gaume et al.*, 2007; *Müller and Haberlandt*, 2016). A reason is that the CSO volume is the most important parameter for practical applications. However, the volume does not explain the complete characteristics of synthetic time series. Consequently, I will consider, both, the CSO volume as well as the duration in the subsequent sections. I will omit the frequency in the following as its characteristics are in most cases very similar compared to the volume, but less robust.

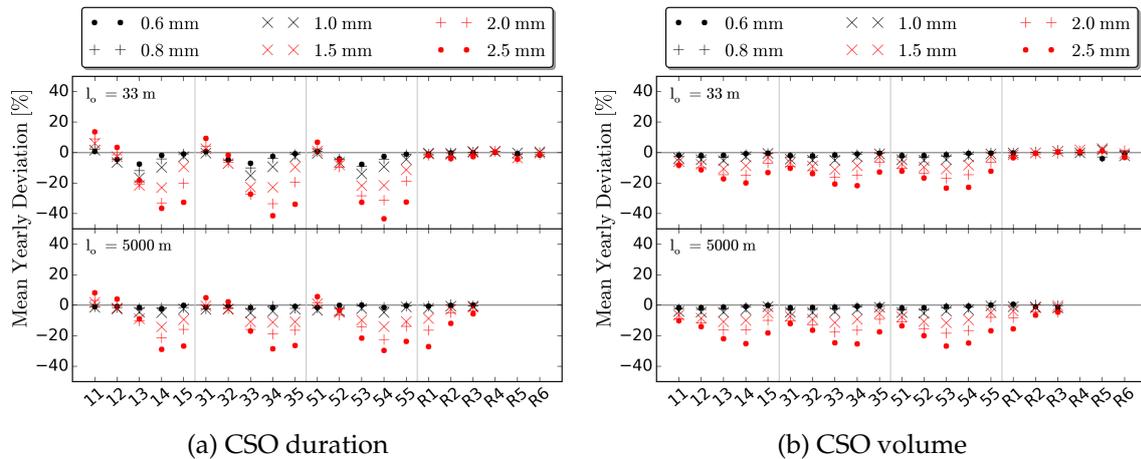


Figure 7.7.: Relative deviation of original and shuffled time series of mean annual CSO statistics. In the shuffled time series, hourly values below a threshold, but larger zero, are randomly mixed. All data correspond to Freiburg (1961 – 1980).

7.2.3. Influence of small values

It is well known that small precipitation events do not contribute to runoff, and thus are unimportant for sewer applications (e.g. *Xanthopoulos, 1990*). However, as *NiedSim* does not optimize event characteristics, but continuous time series at hourly and 5 min aggregation, commonly defined thresholds of events cannot directly be transferred to continuous thresholds. For example, event volumes below 1.5 mm can supposed to be unimportant for sewer simulations. However, an hourly precipitation sum of 1.5 mm could contribute to runoff if the antecedent time period was sufficiently wet. Therefore, I investigate the influence of the temporal structure of small precipitation values here.

As rainfall is highly skewed, a large number of values account for a rather small amount of total precipitation depth. In Fig. 3.2a I demonstrate that around 70 % of all wet 5 min values are smaller than 0.1 mm, but represent only 30 % of the volume. For hourly data the relationship is similar (Fig. 3.2b), where values smaller than 1.0 mm represent approx. 60 % of all values, but explain less than 20 % of the volume. If the influence of these values on the sewer system simulation was very small, less importance could be attached to a large number of precipitation values. Instead, the focus could be on the correct representation of a few, but large values in the *NiedSim* algorithm.

The idea of this analysis is to randomly shuffle precipitation values below a given threshold such that the natural structure is destroyed. By comparing the CSO statistics of the original and the shuffled time series, the influence of the small values on sewer simulations can be observed. The smaller the difference, the less important the values below the threshold are. In order to investigate the structure of the wet periods, I shuffle only wet values such that the dry periods are not affected.

Small hourly values In the first analysis, I shuffle complete hours in 5 min resolution while keeping its measured intra-hourly intensity structure. I mix only hours with a total sum below a threshold $\{0.6 - 2.5 \text{ mm}\}$, but above zero. Figure 7.7 illustrates the relative deviation in CSO statistics compared to the original time series. The influence of small precipitation values depends on the overflow structure and on the length of the overland flow path l_0 . If no storage is available (RR1 – RR6), the deviations are small, which is not surprising because these structures only react on high precipitation intensities.

For structures with storage, both, the overflow volume and duration may be significantly influenced. The results show the high non-linearity of the entire rainfall-runoff-overflow relationship because the actual influence depends on l_0 , q_r , V_s , the applied threshold and the considered CSOs. Nevertheless, a few general characteristics can be observed.

First, the threshold has to be chosen carefully, as hourly precipitation sums $< 1.5 \text{ mm}$, which is a typical threshold for precipitation events, can already lead to more than 20 % deviation in the CSO duration and more than 10 % in the volume. Second, the overflow duration is more effected than the volume which again highlights its large sensitivity to a correct temporal structure. Third, the randomization leads to a systematic underestimation of the CSO statistics. The reason is that small, individual hourly sums are too small to cause overflow on their own. They can only contribute to CSOs in combination with other values such that they form well defined events. The mixing, however, destroys the natural event structure and the distributions at higher temporal scales. Actually, the effect is almost the same as if the values below the threshold were set to zero. The consequence for NiedSim is that a systematic bias could be introduced in sewer simulations, if small precipitation values were not considered in the hourly marginal distribution of synthetic time series.

The fourth and last point is that deviations of more than 40 % can be obtained if hours smaller than 2.5 mm are randomly distributed. This emphasizes the importance of the correct hourly temporal structure for combined sewer systems. Note, these large deviations already occur without changing the wet and dry spell length because only wet hours were shuffled. As the deviations are much larger compared to the influence of the intra-hourly intensity distribution (block rainfall in Fig. 7.6), one can conclude that the hourly optimization of NiedSim has a higher relevance for CSOs compared to the 5 min disaggregation.

Small 5 min values In the second investigation, I shuffle all 5 min values below certain thresholds (Fig. 7.8). The influence is even larger compared to hourly values, which, at first, might be counter intuitive as the individual 5 min values below the thresholds are very small. However, the cumulative hourly precipitation sum with a homogeneous 5 min intensity distribution of 0.2 mm, which is close to the measurement resolution of typical tipping buckets, sums up to 2.4 mm per hour. As shown before, hourly sums of 2.5 mm are important for sewer systems. Therefore, even small 5 min intensities contribute to substantial precipitation at higher aggregations. In other words, the small 5 min intensities play an important role in the definition of large parts of the CDFs at coarser time scales. By the way, the large possible influence of an inaccurate 5 min distribution on higher aggregations is also one reason, why NiedSim does not generate precipitation time series directly on the 5 min aggregation.

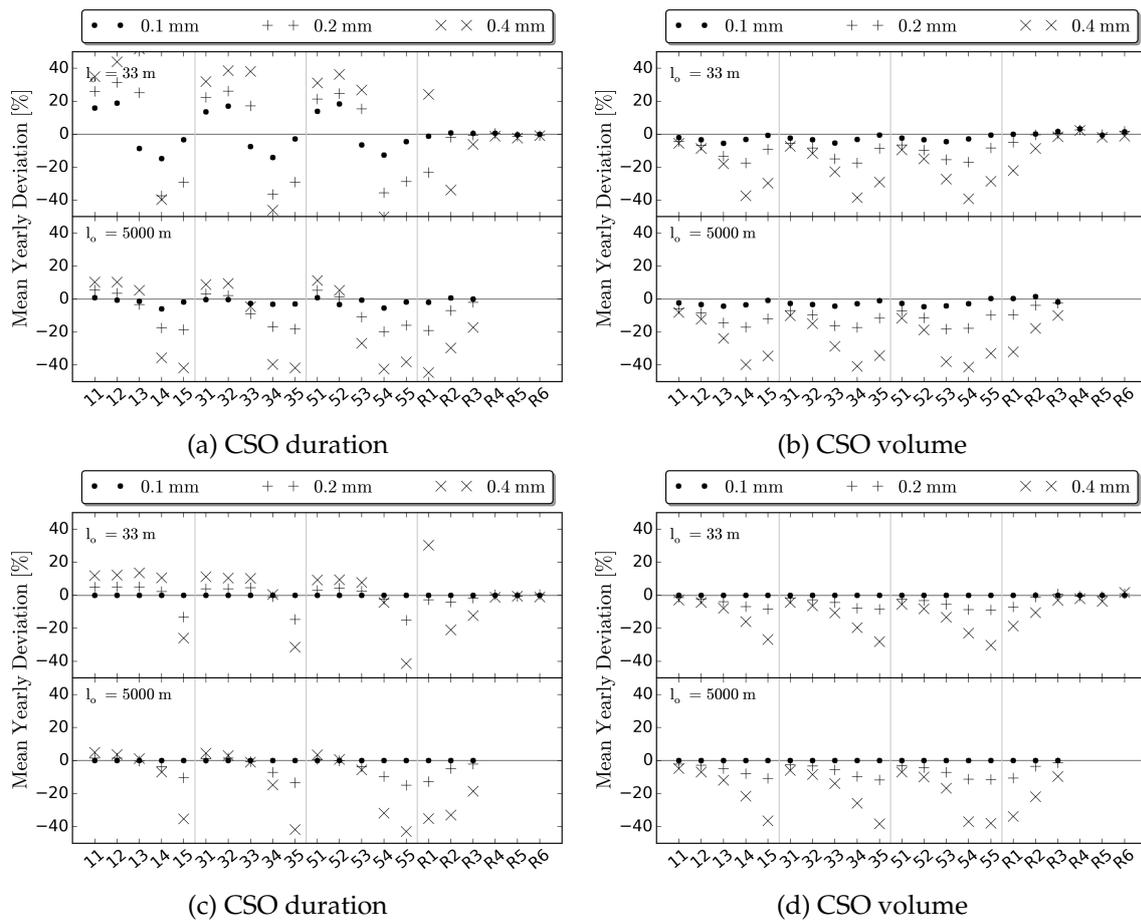


Figure 7.8.: Relative deviation of original and shuffled time series of mean annual CSO statistics. In the shuffled time series, 5 min values below a threshold, but larger zero, were randomly mixed. The data is based on 1961 – 1980 (paper records) in (a, b) and on 1995 – 2005 (tipping buckets) in (c, d). Both sets correspond to data of Freiburg.

It is critical to define a 5 min intensity threshold which is irrelevant for combined sewers because the distribution of small 5 min values highly depends on the measurement device (Fig. 3.2a). A threshold which has minor influence on CSOs, for example, 0.2 mm for tipping buckets (Fig. 7.8c,d), might already have substantial effects for digitized paper records (Fig. 7.8a,b).

Consequently, thresholds at 5 min scales should be used carefully and defined conservatively. In general, it can be concluded that in continuous time series the definition of unimportant values is not as clear as compared to an event-based description of rainfall.

8. Analysis of the hourly optimization of NiedSim

In the previous chapters, I extensively investigated observed rainfall data and developed an indirect validation tool based on an artificial sewer. I will use the findings of both parts in the following to generate and evaluate synthetic precipitation with the model NiedSim. As explained in Chap. 4.4, NiedSim comprises three different, independent parts: the database, the hourly generation and the 5 min disaggregation. The objective of my study is the simulation of precipitation data with the focus on a realistic temporal structure. Consequently, I do not consider the setup of the database including parameter interpolation. Instead, all parameters used here are calculated with measured time series.

The generation of an hourly time series in NiedSim is separated into two parts (Chap. 4.4.2): (a) setup of an initial time series and (b) its temporal optimization. The crucial step of (a) is a realistic description of the hourly distribution function for which different methods exist. For example, NiedSim2.x is based on a parametric approach, whereas *Mosthaf* (2017) illustrates an alternative using non-parametric distributions.

The focus of this study is on (b), the optimization of the temporal structure by rearranging the hourly precipitation values. In order to avoid any influences of (a) on the temporal structure, I do not generate the initial time series. Instead, I destroy the temporal structure of a measured hourly series by shuffling the hours randomly. Afterwards, I use the shuffled hourly series as initial series in the NiedSim 1 h optimization. In this way, the hourly distribution of the measured and optimized series are identical and any deviations in other statistics or in the results of the sewer application can be referred on differences in the temporal structure and, thus, on the optimization of the NiedSim algorithm.

The validation methods used in my study are chosen in regards to the three main application fields of sewer systems: The hydraulic design, the design and simulation of combined sewer systems, and the sewer system management. The precipitation time series must meet different requirements for the individual tasks as I explained in Chap. 5 in detail. To restate, the design of sewers depends on heavy and extreme events. Although I do not explicitly consider the generation of extremes in the optimization, the ability of the NiedSim algorithm to reproduce large events is important. As they can be described well by IDF-curves, I compare IDF-curves of measured and synthetic time series as a first validation criteria. An important point is that I use pure empirical IDF-curves without fitting any probability distribution.

The second validation method I apply here is the artificial sewer introduced in Chap. 6. Its intention is to account for the temporal characteristics of medium spells which are

relevant for design and simulation tasks of combined sewer systems as well as for sewer management.

Before I start with the actual analysis of the hourly optimization, I would like to address the consideration of the extreme values in NiedSim. As explained in Chap. 3.5, extremes are not as robust as other statistics because they are highly vulnerable to sampling errors due to their rare occurrence. Consequently, the uncertainty related to extremes is in general very high. The main objective of NiedSim is not to simulate extreme values, but to generate long continuous time series. However, extremes cannot be simply ignored because they are not only relevant for design purposes, but, at high aggregations, also for combined sewer simulations as shown in Chap. 7.2.2.

NiedSim2.x does not reflect extremes using an explicit extreme value generation approach, but applies the extremes from KOSTRA. The KOSTRA values are used in the initial time series as well as in the objective functions (compare Chap. 4.4). For the following analysis, I removed the extreme values entirely from NiedSim to investigate the ability of the algorithm to reproduce large events without explicitly considering extremes. As a consequence, the resulting time series should be more robust as the uncertainty of extremes does not influence the optimization.

Finally, I would like to mention that I perform all sewer simulations in my study in 5 min resolutions. In order to avoid any influence of the NiedSim disaggregation in the validation of the hourly optimization, the measured and hourly optimized series are both simulated with uniformly distributed intra-hourly intensities (block rainfall).

8.1. Motivation

Large systematic deviations can occur if the synthetic sewer is applied using the precipitation time series of NiedSim2.x. Figure 8.1 shows the relative deviation in the CSO statistics of the optimized hourly time series compared to the measured one. The dark gray bandwidths show the internal variability of 10 different optimization runs. One can clearly observe systematic biases in both statistics that are much larger compared to the confidence intervals of the natural uncertainty (light gray bandwidths). The biases indicate that the temporal structure is not well reconstructed which is the motivation of a systematic analysis of the hourly optimization of NiedSim in this chapter.

Furthermore, the deviations show a high complexity of the rainfall-runoff-processes as the differences vary depending on the regarded CSO statistics and overflow structures. This is a dangerous characteristic because wrong conclusions can be easily drawn from an insufficient indirect validation. For example, if only the volume of B31 with $l_o = 5000$ m was analyzed, the result would pretend an optimized time series with a good temporal structure. Only if all sewer setups are assessed, it turns out that the time series is rather unrealistic. In the first case, the result is just a random fit, where different erroneous precipitation characteristics cancel each other out. Therefore, several different CSO statistics, overflow structures and model setups must be regarded to be able to refer from the sewer output on the general quality of synthetic time series.

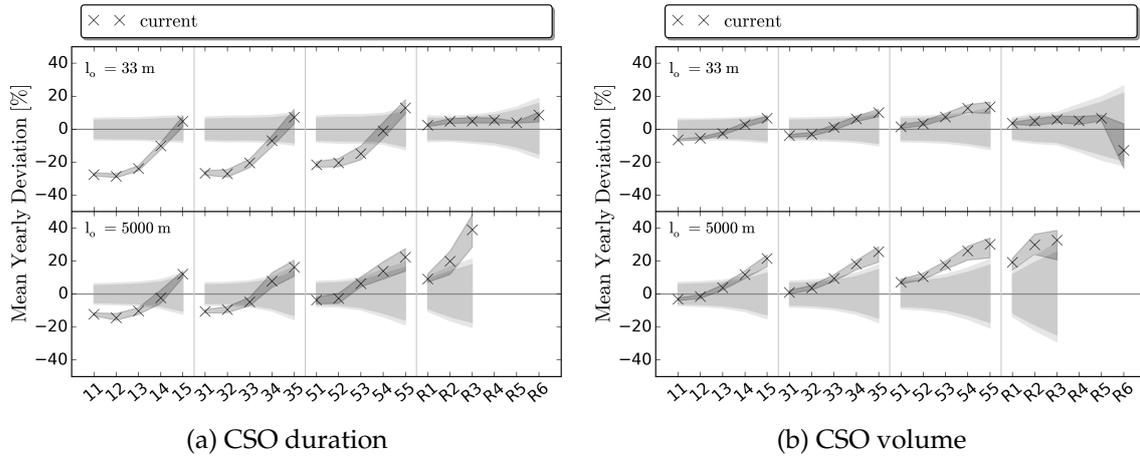


Figure 8.1.: Difference in CSO statistics of measured time series and hourly optimized time series using NiedSim2.x. The dark gray bandwidths illustrate the minimum and maximum of 10 different optimization runs and the crosses indicate their mean. The light gray bandwidths are the 90 / 95 % confidence intervals of 37 years according to Chap. 7.1. All results correspond to Freiburg (1961 – 1980).

Finally, I would like to mention that the internal variability of the NiedSim algorithm of several runs is rather robust as shown in Fig. 8.1. This is a general characteristic of all hourly optimizations. Thus, only the mean result of ten runs is illustrated in the rest of the study without explicitly showing the bandwidth.

8.2. Modifications of the hourly optimization

For the hourly optimization of NiedSim, only the hourly marginal distribution exist as initial boundary condition. Consequently, several parameters are necessary to describe the rainfall characteristics and to reproduce a realistic temporal structure. Nevertheless, the influence of single parameters on the simulation should still be observable to understand their individual influence. Therefore, I have to reduce the parameters of NiedSim2.x (explained in Chap. 4.4) to the most important ones. I use the autocorrelation, the weighted autocorrelation and scaling parameters to reflect the basic temporal structure. Additionally, I take the monthly precipitation sums into account to reflect the seasonal cycle, and the daily exceedance probabilities to support the distribution at a coarse temporal aggregation. I do not use the cross-correlation, the CPs and the daily maxima anymore for the following reasons.

The cross-correlation was introduced to get a spatial consistency if more than one time series is generated. However, a conditional simulation of an entire field using a swapping approach is highly computationally demanding (Beck, 2012). Furthermore, the spatial structure is too complex to be characterized only by the cross-correlation. Therefore, other approaches are pursued for multi-side simulations (Mosthaf, 2017) for which this parameter is irrelevant. Additionally, the focus of my work is on the generation of individual time series for which the cross-correlation has no additional benefit. The reason to disregard the CPs is

Table 8.1.: Overview of the analyzed versions of different hourly optimization configurations. Detailed explanations are given in the respective sections in the text.

Version	Configuration					
	autocorrelations monthly sums daily exc. prob.	scaling param. (used aggregations)	x_{poiss}	n_{poiss}	pI	wet spell freq. (used time period)
a	+	-	-	-	-	-
as	+	sub-daily	-	-	-	-
as+	+	sub-daily + sub-hourly	-	-	-	-
p5	+	-	5 mm	14.2	-	-
p4	+	-	4 mm	23.4	-	-
p3	+	-	3 mm	42.7	-	-
p5x	+	-	5 mm	14.2	89 %	-
p4x	+	-	4 mm	23.4	85 %	-
p3x	+	-	3 mm	42.7	78 %	-
p4xf	+	-	4 mm	23.4	85 %	1961 – 1980
p4xf2	+	-	4 mm	23.4	85 %	2007 – 2012

that they characterize the daily time scales in the way they are implemented in NiedSim2.x. For urban hydrology, however, sub-daily scales are much more relevant than the temporal structure at daily scales. Finally, I neglect the daily maxima because extreme values are not considered in the optimization anymore due to the reasons I explained before.

I would like to mention another difference compared to the version NiedSim2.x that is related to the swapping algorithm. The marginal distributions of the hourly values are systematically different between the seasons, for example, due to heavy convective events that are more likely in summer than in winter. In NiedSim2.x only one marginal distribution is used for the entire year, and thus the swapping of values during the hourly optimization is performed for the whole year as well. The correct seasonal marginals are supposed to be fit indirectly. In my study, however, I use two marginal distributions for winter and summer individually. Consequently, I have to apply an intra-seasonal swapping procedure in the hourly optimization, such that the marginals of each season are not mixed. The seasons are defined according to *Mosthaf* (2017) (Summer: May to August, Winter: September to April). The implementation of an intra-seasonal swapping does not change the optimization, but introduces an additional constraint. If the seasonal marginals are not mixed, the optimization of the target parameters can be improved indirectly as well.

For the following investigations, I use the time period I (196x – 198x, see Chap. 3.2) as standard reference time period because this period is the longest homogeneous time series available for my study. The issues related to high resolution statistics of period I are of minor importance for the hourly optimization. An overview of the different optimization configurations analyzed in the following sections is given in Tab. 8.1.

Version	abs [-]				rel [%]		
	meas	a	as	as+	a	as	as+
$r_1^{(H)}$	0.43	0.44	0.44	0.44	2	2	2
$r_1^{(2H)}$	0.36	0.37	0.37	0.37	3	3	3
$r_1^{(3H)}$	0.33	0.32	0.32	0.32	-3	-3	-3
$r_1^{(6H)}$	0.27	0.27	0.27	0.27	0	0	0
$r_1^{(12H)}$	0.22	0.21	0.22	0.22	0	-5	0
$r_1^{(D)}$	0.16	0.16	0.16	0.16	0	0	0
b_1^{summer}	0.61	0.66	0.65	0.72	9	7	18
b_1^{winter}	0.64	0.68	0.68	0.72	6	6	11
b_2^{summer}	1.04	1.10	1.09	1.23	6	4	18
b_2^{winter}	1.25	1.25	1.30	1.43	0	6	14
b_3^{summer}	1.32	1.41	1.40	1.61	7	6	22
b_3^{winter}	1.86	1.74	1.88	2.14	-6	1	15

Table 8.2.: Overview of autocorrelation r and scaling parameters b (of sub-daily moments) of measured time series (*meas*) and optimized time series (*a*, *as*, *as+*) using different versions of NiedSim (see Tab. 8.1). All data refer to Freiburg, 1961 – 1980.

8.2.1. Autocorrelation and scaling parameters

The autocorrelation and scaling parameters are used to characterize the basic temporal structure in NiedSim. In this section, I analyze if these parameters are sufficient for a realistic time series or if the issues observed in Fig. 8.1 can be related to these parameters. For this purpose, I define three different objective functions. Then I use the objective functions to optimize the randomly shuffled hourly measured time series. The first objective function (version *a*) includes the autocorrelation, weighted autocorrelation, monthly precipitation sum and daily exceedance probabilities. In the version *as* and *as+*, the scaling parameters are additionally incorporated as target parameters. The moments I used for the calculation of the scaling parameters are different in both versions. In version *as* I take only sub-daily¹ moments of measured series, whereas in version *as+* I use all scales from 5 min to daily aggregation. The latter one is applied in NiedSim2.x. The comparison of *as* and *as+* can show the impact of scaling breaks on the rainfall generation. If no scaling breaks occurred, it would not matter if sub-hourly scales were used or not. However, due to scaling breaks illustrated in Chap. 3.4.1, the scaling parameters highly differ depending on the used aggregations.

At first, the statistics of the simulated time series are directly compared to the statistics of the measured time series to investigate if the algorithm is able to optimize the parameters at all. Table 8.2 illustrates the autocorrelation and scaling parameters (of sub-daily moments). The table shows that the autocorrelation and scaling parameters can already be optimized well, even if the scaling parameters are not part of the objective function (version *a*). The

¹In the context of scaling parameters, sub-daily is defined as time scales from hourly to daily aggregation and does not include higher resolutions.

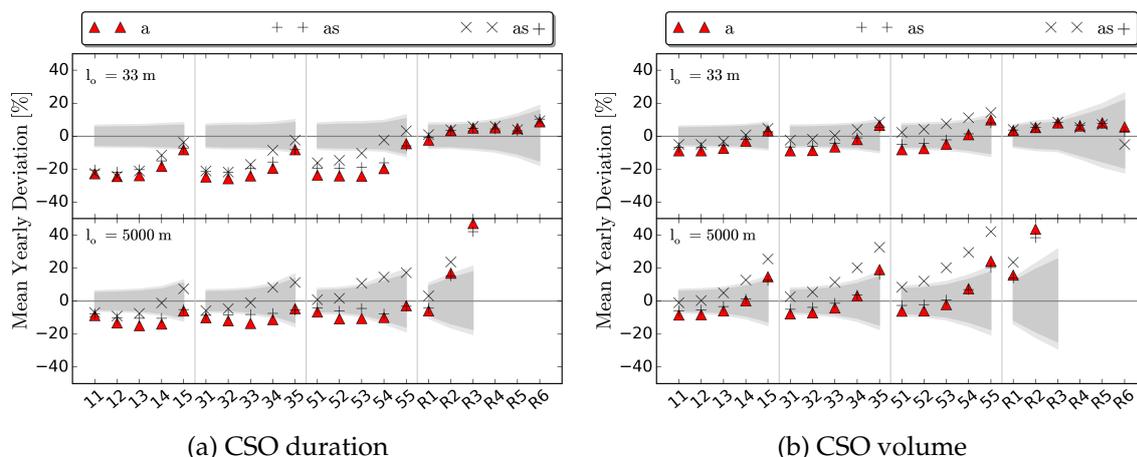


Figure 8.2.: Differences in CSO statistics of measured time series and hourly optimized time series using NiedSim (mean of 10 runs). The versions are defined according to Tab. 8.1. The light gray bandwidths are the 90 / 95 % confidence intervals of 37 years according to Chap. 7.1. All results correspond to Freiburg (1961 – 1980).

additional optimization of the scaling parameters (version *as*) can not significantly improve the scaling parameters. The large deviations of version *as+* are related to the larger target scaling parameters due to the usage of sub-daily time scales. As shown in Chap. 3.4, the sub-hourly time scales lead to larger scaling parameters, which are not only caused by scaling breaks, but also by different measurement devices. Here, the effect is especially large because the target parameters are calculated with data from 1961 – 1980 (digitized paper records, period I). This time period is characterized by a rather homogeneous sub-hourly structure with large scaling parameters that are not representative for sub-daily time scales.

No differences in the optimization of the autocorrelation between the three versions can be observed and only small deviations occur for the scaling parameters. That is, both parameters can be optimized simultaneously, although the parameters are not independent of each other.

As the deviations of version *a* and *as* are comparably small, I apply the indirect validation using the sewer system to investigate the relevance of an explicit optimization on scaling parameters. As Fig. 8.2 illustrates, slight deviations in the scaling statistics of version *a* and *as* have only a small effect on the application. However, as soon as the scaling parameters are substantially increased (version *as+*), the influence on the temporal structure is significant and leads to large overestimations of the CSO volume. The issues of NiedSim2.x with respect to CSO statistics shown in Fig. 8.1 are similar to the results of version *as+*. Therefore, they can, at least partially, be referred to the changing characteristics of the scaling behavior from scaling breaks and from different measurement devices.

Some additional aspects can be shown if the deviations in the partial series are considered (Fig. 8.3). The red lines in the figures indicate the values which would be used if the partial series were obtained with the definition according to KOSTRA (i.e. 2.7 values/year on average, see Chap. 2.2.3). However, to be able to link the partial series with the results of the

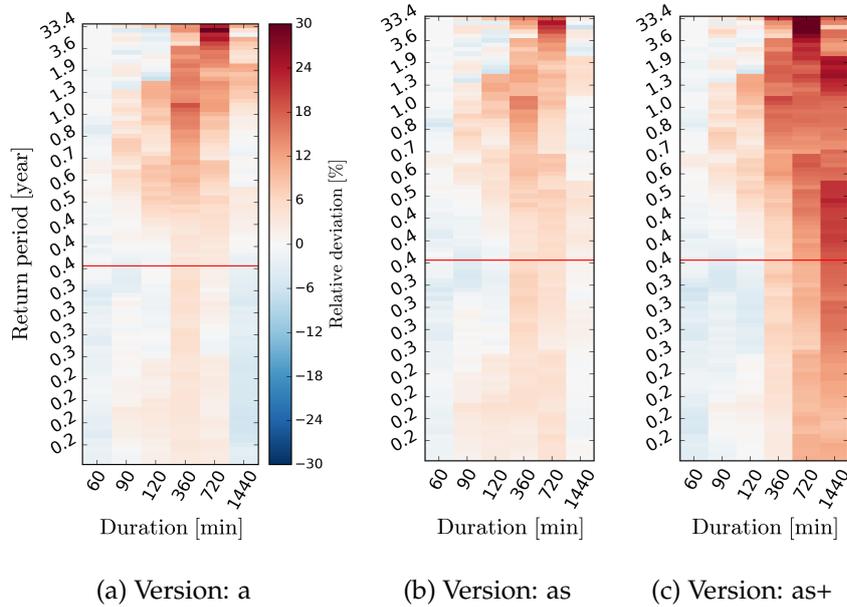


Figure 8.3.: Relative deviation of partial series of different versions of hourly optimized time series (see Tab. 8.1) from the measured time series. Return periods above the red line are usually considered for extreme value analysis. All data refer to Freiburg, 1961 – 1980.

CSO statistics, I extend the partial series to smaller values reflecting lower return periods (5 values/year on average). The large influence of the target scaling parameters on the partial series is obvious. The larger the scaling parameters in the objective function, the more the partial series are overestimated at higher time scales.

I would like to highlight the relevance of both validation techniques again. If only the partial series of version *a* or *as* were analyzed, one would conclude an overestimation of extreme values by the synthetic time series. This conclusion, however, cannot be transferred to the behavior of the time series in combined sewer simulations, where the CSO statistics are underestimated for many overflow structures. Even if more values are included as compared to KOSTRA (below the red line), the systematic underestimation cannot be observed in the partial series.

In conclusion, the usage of scaling parameters in the objective function leads to comparable time series as if only autocorrelations are optimized. However, wrong scaling parameters used as targets can considerably influence the optimization. Additionally, as scaling parameters of higher moments are rather vulnerable to outliers, I do not further use the scaling parameters in NiedSim to make the hourly optimization more robust.

Nevertheless, systematic biases in the CSO statistics as well as in the partial series can be noticed in version *a*. They indicate further problems of the optimization algorithm apart from the issues with scaling parameters. On the one hand, large precipitation values tend to be overestimated (Fig. 8.3) and lead to an overestimation of the CSO volume (Fig. 8.2) for structures with larger volumes (e.g. 55) or large specific discharges (e.g. R1). On the

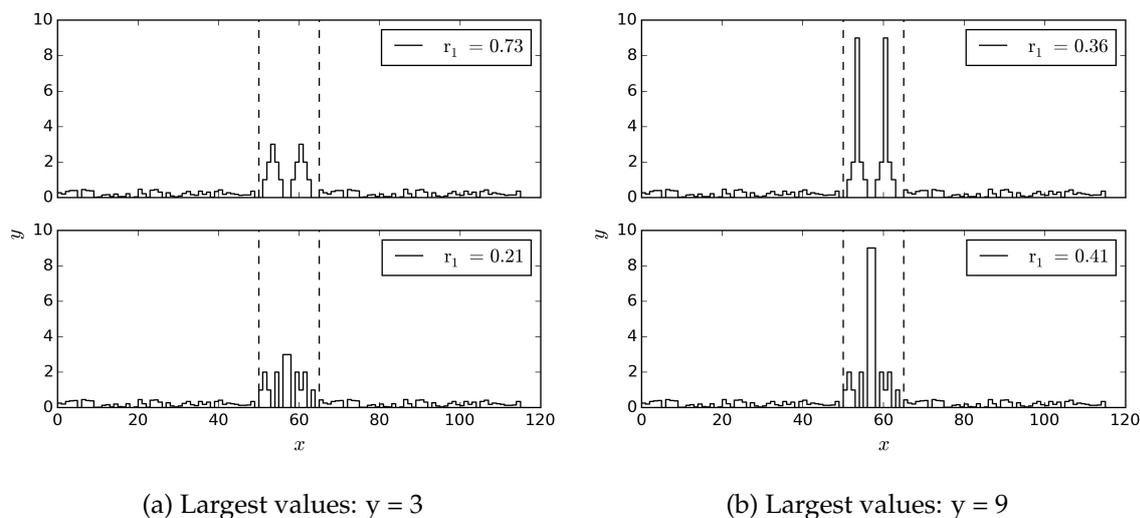


Figure 8.4.: Time series and autocorrelation r_1 for a fictive spell embedded in small random values. In the upper case, the spell is structured with two peaks around the largest values y . In the lower case, both y are connected in the middle of the spell and surrounded by values with a random structure.

other hand, for structures (e.g. 11, 31) or statistics (duration), where smaller values are relevant, the CSO statistics are underestimated. This suggests that the natural precipitation time series exhibits a higher complexity than what can be reflected by autocorrelation, scaling parameters, monthly precipitation sums and daily exceedance probabilities.

8.2.2. Positioning of large values in a Poisson process

A major issue of the optimization discussed in the previous chapter is the overestimation of large precipitation spells at durations larger than one hour (Fig. 8.3). As the problems occur even if no scaling parameters are considered, they must be caused by the autocorrelation. Therefore, some additional thoughts about the optimization of the autocorrelation are considered first.

8.2.2.1. Initial considerations

The illustrated issues result from the autocorrelation that is a non-reversible function. I use the term non-reversible to describe the characteristic that several precipitation time series can have the same autocorrelation values, but may exhibit very different temporal structures. This characteristic can be illustrated with an artificial example.

Consider a precipitation spell $[0,0,1,2,y,2,1,0,0,1,2,y,2,1,0,0]$ that shows two symmetrical peaks around the maximum precipitation y , where $y = 3$ (Fig. 8.4a, top panel). This spell is embedded in 100 uniformly distributed random values between $[0, 1]$ leading to skewed

marginals. In Fig. 8.4a (bottom panel), the same marginals are shown, but the spell is randomized except for the two largest values y . Both y values are connected and centered in the middle of the spell. As expected, the autocorrelation of the second spell is reduced ($r_1 = 0.21$) compared to the first one ($r_1 = 0.73$) due to the randomization of the spell.

Now, consider another data set with even more skewed marginals, that is $y = 9$, without changing the other values. The autocorrelation of the well structured spell reduces to $r_1 = 0.36$ (Fig. 8.4b, top panel), whereas for the unstructured spell r_1 increases to 0.41 (Fig. 8.4b, bottom panel). Therefore, the position of the largest values y within the time series affect the autocorrelation more than the structure of smaller values. As the largest values are connected in the unstructured spell, but are unconnected in the structured spell, the autocorrelation in the unstructured spell is higher, even if most values are randomly located.

From a perspective of the NiedSim optimization algorithm, it is much easier to optimize a high autocorrelation with an unstructured spell, where only a few large values are connected and need to be swapped correctly, than in a structured spell. Consequently, the autocorrelation in NiedSim tends to cluster large values which leads to an overestimation at higher time scales. On the other hand, the structure of medium and small values is rarely influenced by the optimization and, thus, remains rather random.

From a meteorological point of view, there exist at least two major processes that generate precipitation: First, long lasting stratiform rainfall of medium intensity, and second, short, but heavy convective precipitation. The statistics of both forms are significantly different. In NiedSim2.x these differences are only reflected by differentiating between the summer and the winter half-year. The underlying assumption is that both seasons behave differently because in summer more convective events occur than in winter. Consequently, the weighted autocorrelations in NiedSim2.x should trigger seasonal characteristics reflecting the differences between summer and winter rainfall. However, stratiform and convective events may occur in both seasons, but with different probabilities, and thus the actual seasonal statistics are influenced by both phenomena. The resulting seasonal autocorrelations are affected by a mixture of different precipitation events, but reversely, differences in the autocorrelation do not automatically account for different precipitation processes during the optimization. As in NiedSim2.x, no process is involved that reflects the differences between convective and stratiform rainfall events, the optimization routine optimizes the time series in its easiest way, that is, by clustering large precipitation values. However, these events are neither realistic stratiform nor convective events, but simply reflect the target autocorrelation values.

The idea to improve the NiedSim algorithm is to introduce an additional process that avoids the clustering of large values and supports the generation of short convective events. For this purpose, the positions of large values are predefined in the initial time series of NiedSim and are not allowed to be swapped afterwards in the hourly optimization. As in this case, the target autocorrelation cannot be obtained by adjusting a few large values anymore, the algorithm must optimize the bulk of medium values reflecting the characteristics of stratiform rainfall. The positions of the large values are determined using a stochastic process, the Poisson process, which has been successfully used to trigger precipitation cells in Point Process rainfall models (see Chap. 4.3).

8.2.2.2. General Poisson process

A Poisson process is a counting process (*Marshall and Olkin, 2007*), where the probability that a random variable X occurs k times in a time interval is Poisson distributed with

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad (8.1)$$

where λ is the rate parameter. The expected value of the Poisson distribution is equal to the rate parameter λ and describes the mean number of occurrences within the defined interval. In a Poisson process, the inter-arrival time t between two occurrences is exponentially distributed with the CDF

$$F(t) = 1 - e^{-\lambda t}, \quad t \geq 0, \quad (8.2)$$

The expected value of the exponential distribution is simply the reciprocal of the occurrence frequency, that is $1/\lambda$, and indicates the mean inter-arrival time between two occurrences.

In this study, I use the Poisson process to define the location of a given number of largest precipitation values n_{poiss} within a month. For this purpose, I have to apply several modifications to the general Poisson process.

First, the parameter λ can not supposed to be constant because the inter-arrival time between two large values differs within the year. In summer months, a higher number of large values occur than in winter months, respectively the inter-arrival times are smaller. Additionally, the probability of large values occurring in a certain month depends on the monthly precipitation sum. Obviously, if a month is dry, no (large) values can occur, and thus none should be placed with the Poisson process. In order to reflect these characteristics, I define a variable parameter λ that depends on the time within the year as well as on the monthly precipitation sum. The following section illustrates how I fit λ to fulfill these requirements.

8.2.2.3. Fitting of the parameter lambda

The fitting procedure of λ contains two steps: First, the description of the dependence between the number of large hourly rainfall values per month and monthly precipitation, and second, the characterization of this dependence within the annual cycle.

The initial step is the selection of the number of largest values per year n_{poiss} , which are intended to be set in the Poisson process. n_{poiss} is defined indirectly by using values larger than a selected precipitation threshold x_{poiss} (e.g. $x_{poiss} > 4$ mm corresponding to $n_{poiss} \approx 23$ values). If the number of n_{poiss} is defined directly, it is possible that rather small values are taken into account for which the relationship between monthly precipitation and n_{poiss} is less pronounced. The result would be a less robust fitting of λ . For each of the n_{poiss} values, the respective month is obtained in which it occurs. The number of occurrences per month is counted and leads to a set of monthly frequencies $n_{poiss,m}$ where m denotes the month. Additionally, the respective monthly precipitation sums H_m are calculated in which the values occur.

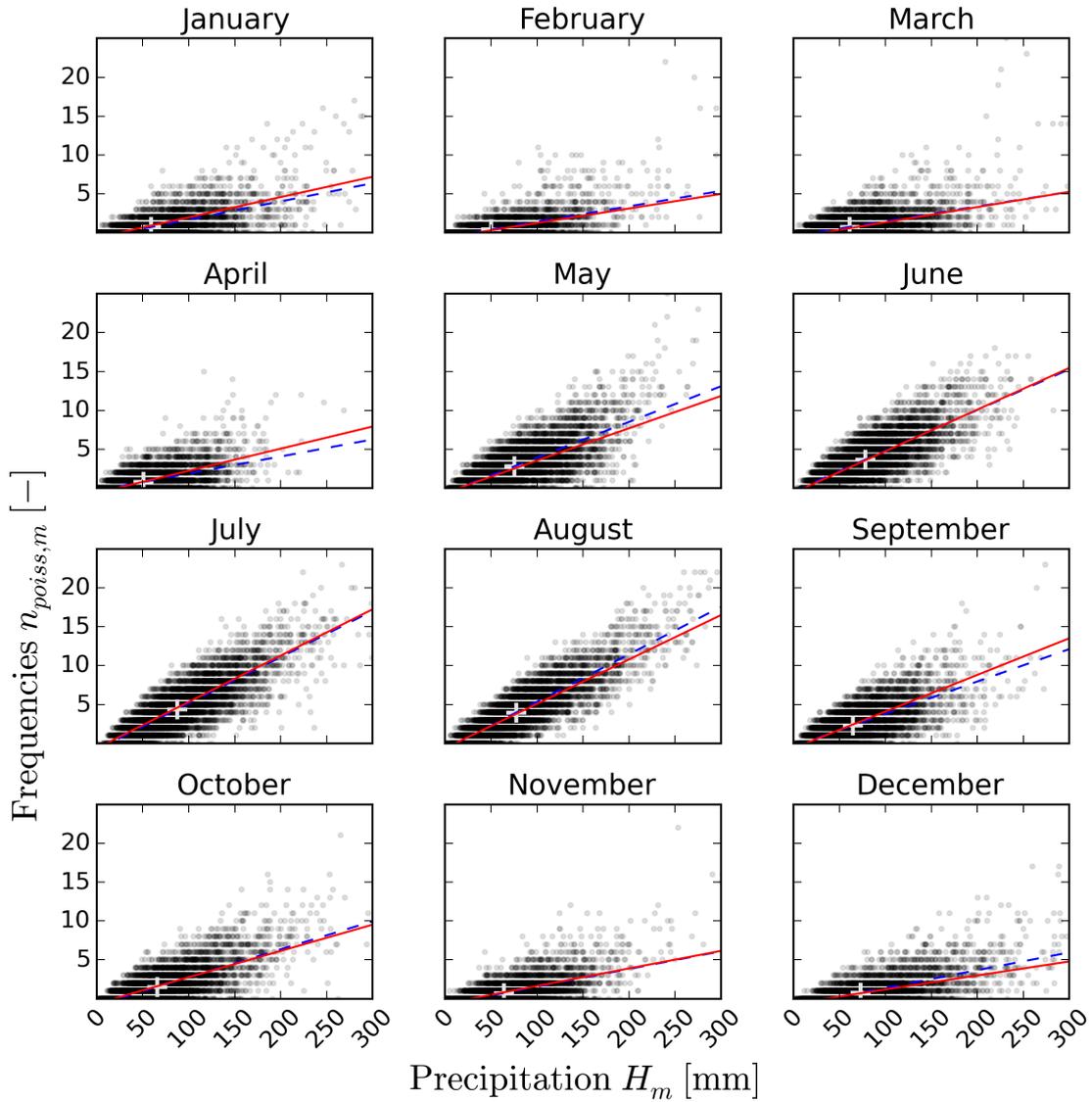


Figure 8.5.: Frequencies of values ($x_{poiiss} > 4$ mm) and monthly precipitation sums of 1092 gauges in south Germany. The cross dedicates the mean frequency $\bar{n}_{poiiss,m}$ and mean precipitation sum \bar{H}_m . The dashed line reflects a linear regression of the measured data. The red line is the fitted line with a mean monthly intersection and a slope reflecting the annual cycle with a sine curve.

The frequencies of occurrence and the precipitation sum can be plotted in a scatter plot for each month individually to observe their relationship. If only one pluviometer is considered, no clear structure can be observed due to rather small sample sizes per month, but large variances of the monthly precipitation sums. Consequently, I use the hourly data of south Germany of 1092 stations, where each pluviometer has available data of at least five years with a minimum of 8500 hours per year ($\approx 97\%$). I combine the data of all years and all stations into one set of frequencies and monthly sums with a total length of 7515 years. The procedure assumes a general dependence between $n_{poiss,m}$ and H_m independent of the location – an assumption, I will verify later in this study.

Figure 8.5 illustrates the dependence between $n_{poiss,m}$ and H_m for $x_{poiss} > 4$ mm. The white crosses indicate the average frequencies $\bar{n}_{poiss,m}$ and the mean precipitation sums \bar{H}_m of each month. As expected, one can clearly notice that large precipitation values occur more frequently in summer months (caused by convective events), and that the frequency increases with larger monthly precipitation sums. For each month, I perform a linear regression of all 7515 years leading to the increasing, blue dashed lines in Fig. 8.5. The parameters (intersection a_m and slope b_m) of the regression enable to define the number of values to be set with the Poisson process (λ_m) within a month m for a given monthly precipitation sum H_m according to

$$\lambda_m = n_{poiss,m} = a_m + b_m * H_m \quad (8.3)$$

The changing slope of the regression line within the month accounts for the annual cycle. In order to reduce the discrete jumps of the slopes between months, I fit a continuous sine curve to the slopes with

$$b(t_i) = a * \sin(t_i - c) + d, \quad (8.4)$$

where $b(t_i)$ defines the slope at time t_i and a , b , c and d are parameters of the general sine function. The supporting points for the fitting of the sine function are the 12 monthly slopes b_m . The result of the fitted sine curve is shown in Fig. 8.6. The intersections of all months (same figure) do not show an annual cycle, but are rather constant. Therefore, I use the mean intersection \bar{a} of all months for the continuous representation of the dependence, and Eq. 8.3 results in

$$\lambda(t_i) = \bar{a} + b(t_i) * H_m \quad (8.5)$$

where $\lambda(t_i)$ is the number of values set with the Poisson process at an hourly time step i within the year. The regression lines in the fitted, continuous form are shown in Fig. 8.5 as red lines. Only minor deviations from the discrete monthly regressions can be observed.

I also tested an alternative fitting procedure of λ using a least-squares optimization of a 3d-surface (Fig. A.1) that I finally did not use. In summer months the number of Poisson values is greater than zero even if the monthly precipitation is zero. The resulting potential errors in the optimization do not occur in the linear regression approach because all months have a negative intersection.

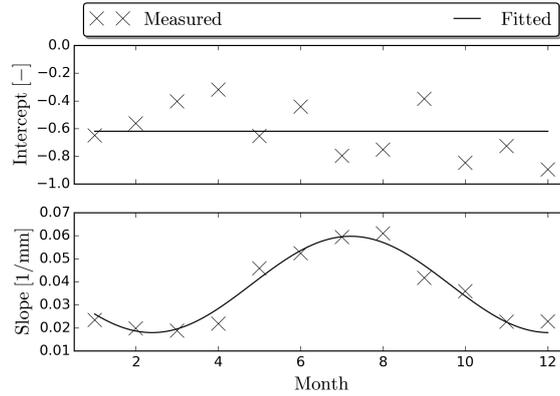


Figure 8.6.: Regression parameters intercept (a_m) and slope (b_m) of measured data and fitted sine curve representing the continuous slope $b(t_i)$. The figure shows the data for $x_{poiss} > 4$ mm.

Finally, I verify the assumption from the beginning, which stated a global dependence structure of the occurrence of large values. At first, I illustrate the independence of the elevation. For this purpose, I plot the occurrence frequencies per month of all 7515 years against the elevations of the respective gauges. Only a slight tendency of an increasing number of occurrences with higher elevations can be noticed (Fig. A.2). A linear regression leads to correlation coefficients of around 0.25 in summer months. This weak dependence is not caused by the elevation, but by a secondary variable, namely the monthly precipitation sum, which is generally larger for higher elevations. If I standardize the Poisson values by their monthly precipitation sums $n_{poiss,m}/H_m$, no correlation can be detected anymore (Fig. A.3).

Even if the occurrence of large values is not correlated with elevation, a spatial dependence structure might exist. A possibility to analyze such a relationship are empirical variograms (Wackernagel, 1995). A variogram $\gamma(h)$ is a function that relates the distance h between two points in space with the squared differences of a variable of interest between both points. The higher the dependence, the lower the differences between the two points, and thus the lower $\gamma(h)$ is. In case of spatial dependence, the variogram rises from $\gamma(h=0) = 0$ with increasing distance to a maximum value (sill) which can be equal to the variance of the underlying process. In Fig. A.4 the variogram of the occurrence frequencies is plotted. Their dependence structure is very weak because even for short distances, the variogram $\gamma(h \rightarrow 0)$ is already high and remains almost constant with increasing h .

In conclusion, the underlying process, which is relevant for the occurrences of large values per month, does neither show a spatial dependence, nor a correlation with the elevation. It depends only on the monthly precipitation sum. Consequently, one can combine all available rainfall observations to estimate λ most robustly.

Finally, I would like to mention that the dependence analyzed above is not only valid for values $x_{poiss} > 4$ mm, but also for other thresholds ($x_{poiss} > 3$ mm and $x_{poiss} > 5$ mm), although not explicitly shown here.

8.2.2.4. Modification for dependent large hourly values

In the procedure explained so far, all values above a certain threshold are set with a Poisson process. However, as the average inter-arrival time is in the range of days, it is very unlikely that more than one hourly Poisson value is part of the same precipitation event. In measured series several large values may occur in the same event. Consequently, the partial series would be underestimated, in particular, if n_{poiss} was high. This issue can also be explained from another point of view. An important assumption of a Poisson process is the independence of the individual occurrences. For large hourly precipitation values, however, this assumption cannot be hold. Therefore, I modify the Poisson process further in this section.

The concept here is to enable the swapping of at least some of the largest values in the optimization routine of NiedSim. This is achieved by the pre-positioning of the same number of values in the initial time series, but these values need not to be the largest ones anymore. That is, the position of the values are obtained with the Poisson process as defined above, but instead of using all n_{poiss} largest values, an extended pool of potential values is defined according to

$$n_{poiss}^+ = n_{poiss}(2 - p_I), \quad (8.6)$$

where n_{poiss}^+ is the increased number of potential Poisson values and p_I is defined as the probability of n_{poiss} being independent. From the extended pool of values n_{poiss}^+ , n_{poiss} values are drawn randomly and positioned with the Poisson process.

The probability of independence p_I is estimated from measured data. The values are supposed to be independent if not more than one Poisson value occurs in the same precipitation spell separated by at least one dry hour.

8.2.2.5. Implementation in NiedSim

I implement the Poisson process in NiedSim as follows. Before the hourly optimization is started, two Poisson processes are initiated to define the position of the large values. Both processes begin in June; one goes in forward direction from June to December and the other one is reversed from June to January. The reason for two processes is that entire months can be skipped if λ is very low, and thus the inter-arrival time can be very high. For summer months the probability of having a very large inter-arrival time is lower. The inter-arrival time between two large events t_i and t_{i+1} can be calculated using the inverse of the exponential distribution (Eq. 8.2) with λ being $\lambda(t_i)$ obtained by Eq. (8.5).

However, even in summer months, a very small λ could be drawn, such that no value would be set throughout the rest of the year. Therefore, the inter-arrival time must be limited which could be done using a truncated exponential distribution. Here, a maximum waiting time can be set in which at least one value is drawn. Unfortunately, the maximum waiting time is difficult to define. On the one hand, the limit should be small enough to avoid a skipping of entire (potentially wet) months. On the other hand, in dry winter months

the number of Poisson values might be zero for several months in a row, which cannot be reflected if a maximum waiting time is defined. Therefore, I applied a different method using the regular exponential distribution, but in a modified way. If the inter-arrival time is larger than thirty days, the rest of the current month is skipped without setting a value and a new process is triggered on the first day of the next month. I also apply the same procedure if $\lambda < 0$ is drawn.

The question remains, how many large values n_{poiss} should be positioned with the Poisson process to improve the NiedSim optimization. For this purpose, I simulate time series with three different versions ($p5$, $p4$, $p3$) using $x_{poiss} > \{5.0, 4.0, 3.0 \text{ mm}\}$ which reflect a mean number of Poisson values $\bar{n}_{poiss} = \{14.2, 23.4, 42.7\}$ respectively. Additionally, I generate time series with three versions ($p5x$, $p4x$, $p3x$) with the same Poisson processes, but with the dependence criterion explained in Chap. 8.2.2.4. Remark: Compared to \bar{n}_{poiss} , the actual n_{poiss} of a simulation can differ from year to year as well as between different stations because of its relationship to the variable monthly precipitation sum.

I implement the Poisson process to prevent the clustering of large hourly precipitation values. If the partial series in Fig. 8.7 are considered, the overestimation of precipitation values at higher time scales in the partial series disappears using a Poisson process ($p5 - p3$) compared to the algorithm without the Poisson Process (a). The more large values are positioned with the Poisson process, the more the partial series for high return periods are underestimated. This is not surprising as the large values are scattered within the time series and cannot cluster anymore to form extreme events. If the optimization is able to swap at least some of the large values in the modified Poisson process ($p5x - p3x$), the underestimation is reduced. Furthermore, the modified Poisson process is also more robust with respect to the limit x_{poiss} because of a self-controlling characteristic. The more values are set in the Poisson process, the lower p_I will be, and thus the more likely it is that some of the large values are not pre-positioned in the time series. In the case of Freiburg this effect is rather small, but for other locations like Kempten (Fig. A.5), the modified Poisson process yields better and more robust results.

Not only the partial series are affected, but also the entire hourly temporal structure is improved as shown by the reduced deviations in the CSO statistics in Fig. 8.8. The large values, that are not able to cluster anymore, trigger additional, substantial precipitation events that are able to cause individual CSOs, instead of generating only a lower number of very large precipitation spells.

The question about the parameter choice remains, that is, how the threshold x_{poiss} should be chosen. The answer is ambiguous because it depends on the regarded application and location. By applying a threshold $x_{poiss} = 3 \text{ mm}$, too many values are pre-positioned because not only the partial series is underestimated, but also the CSO volume and duration for structures with a high specific discharge (35, 55). By setting a few values ($x_{poiss} = 5 \text{ mm}$), the partial series can still be overestimated (e.g. Kempten) and the positive effect of the Poisson Process on the CSO duration can be rather low. Therefore, I suppose a threshold $x_{poiss} = 4 \text{ mm}$ to be a good parameter for the Poisson process. An optimal parameter is difficult to find due to the complexity of the system. However, the parameter selection is not sensitive because the results are very robust for different x_{poiss} due to the self-controlling

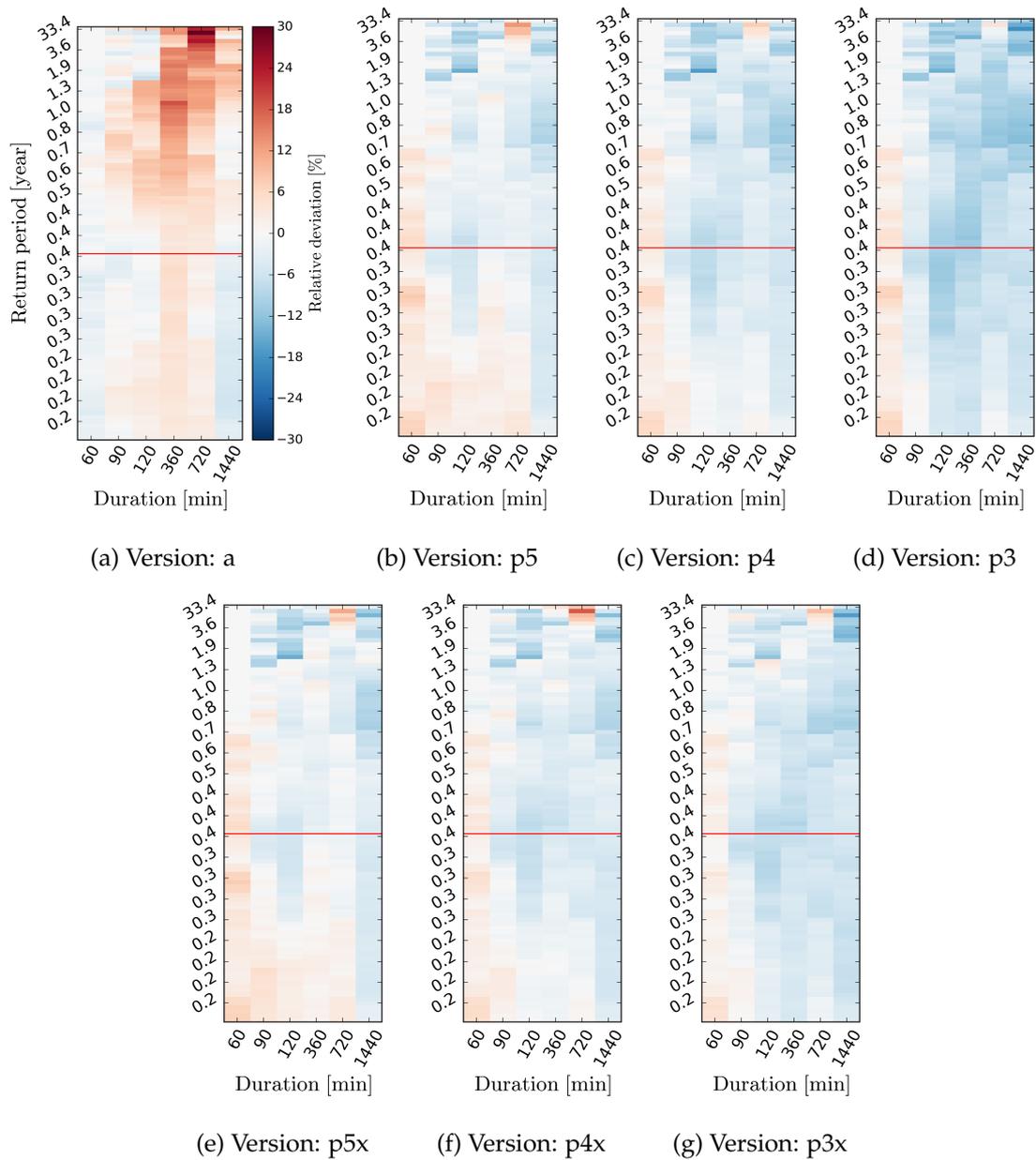


Figure 8.7.: Relative deviation of partial series of different versions of hourly optimized time series (see Tab. 8.1) from the original time series. Return periods above the red line are usually considered for extreme value analysis. All data refer to Freiburg (1961 – 1980).

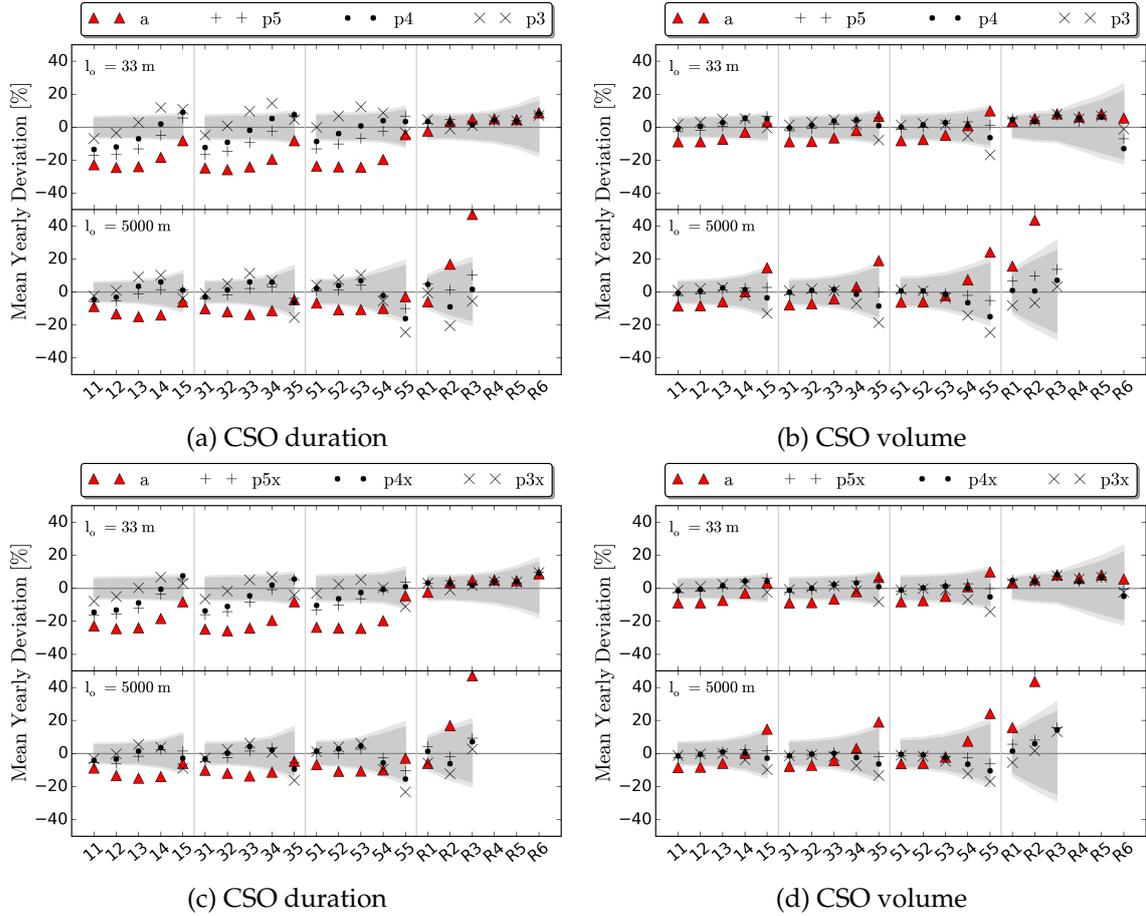


Figure 8.8.: Differences in CSO statistics of measured time series and hourly optimized time series using NiedSim (mean of 10 runs). The versions are defined according to Tab. 8.1. The light gray bandwidths are the 90 / 95 % confidence intervals of 37 years according to Chap. 7.1. All results correspond to Freiburg (1961 – 1980).

mechanism of the modified Poisson process as I explained before.

In conclusion, the Poisson process is a necessary part of the NiedSim algorithm that improves the temporal structure of the time series. The partial series as well as the CSO volume can be well reproduced. The CSO duration, however, which is most sensitive to medium and small precipitation values, still shows large systematic biases.

8.2.3. Including wet spell durations

The optimization of the precipitation time series is still insufficient with respect to some characteristics of the temporal structure even after the Poisson process is introduced. A property not explicitly considered so far is the intermittency, that is, the change between wet and dry periods. Except for the daily exceedance probabilities (0 mm/day), no other parameter is included in the optimization which is somehow related to intermittency. The

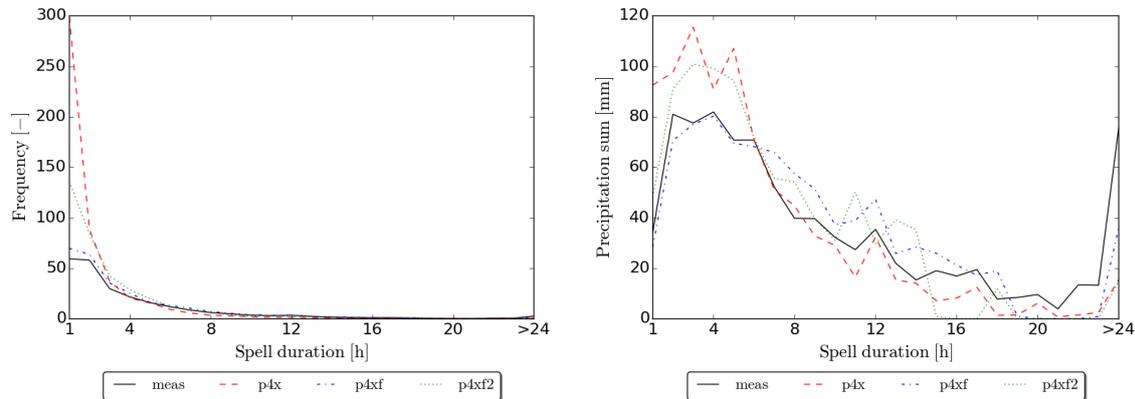


Figure 8.9.: Histogram of wet spell durations (left panel), where the spells are separated by at least one dry hour ($t_{s_e} = 1$ h) as well as the precipitation sum contributed by each duration class (right panel). The four different cases are measured data and the versions according to Tab. 8.1. All data refer to Freiburg (1961 – 1980).

intermittency is often defined with p_0 at different timescales. Another way of reflecting the intermittency indirectly is the description of wet and dry precipitation events. However, such a procedure always comes along with the necessity to separate individual events, which is an ambiguous task as I explained in Chap. 2.2.2. The way I consider the intermittency evades a proper event definition criteria by simply defining precipitation spells as consecutive values larger than zero separated by at least one zero value. Obviously, such a simple criterion is not an appropriate definition of a precipitation event. However, it can serve as additional constraint in NiedSim that can help to improve the optimization as I will show in the following.

In Fig. 8.9 (left panel) the histogram of spell durations of measured and optimized time series are plotted. The best optimization according to the previous section (version $p4x$) shows too many very short individual spells (up to six times more spells). The spell frequency is less important for applications because it only counts number of spells of a certain length. Therefore, Fig. 8.9 (right panel) illustrates the precipitation amount that is contributed by the spells of a classified length. The rainfall sum of short spells is also overestimated, but only twice as high. Therefore, the majority of short spells must be rather small with respect to the rainfall depth. On the other hand, the precipitation sums of longer spells (> 8 h) are underestimated. Both characteristics can influence the sewer simulations. If small values cluster together and form individual spells, they cannot contribute to CSOs, which would be the case if they were part of a larger spell (see Chap. 7.2.3).

In order to reduce the scattering of individual hours, the distribution of wet spell durations separated by one dry hour is introduced as target parameter in the hourly optimization routine. That is, the difference of the discrete probability density function (PDF) of the optimized time series and the empirical target PDF obtained from measured time series is minimized. The NiedSim optimization is computationally very expensive as the objective function of several hundred thousands of swaps has to be evaluated. The Simulated An-

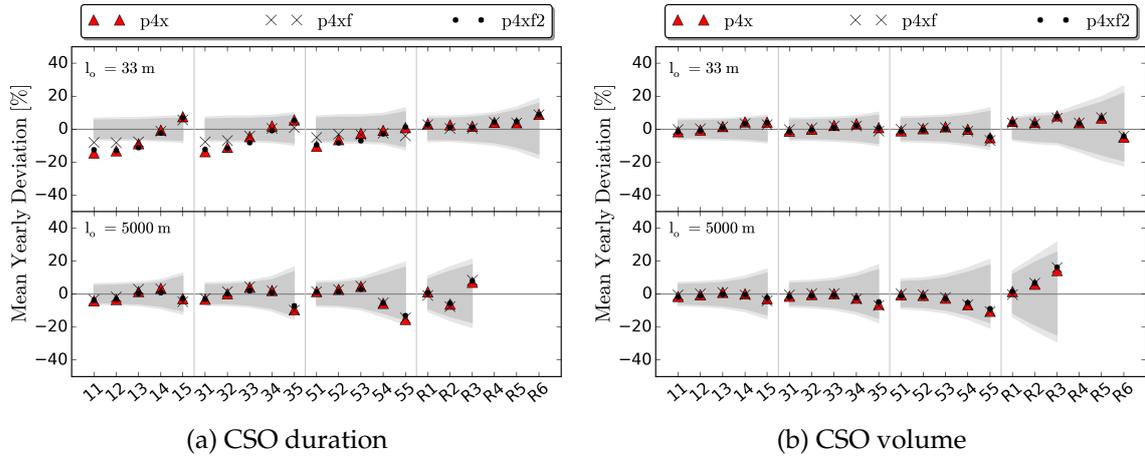


Figure 8.10.: Differences in CSO statistics of measured time series and hourly optimized time series using NiedSim (mean of 10 runs). The four different cases are measured data and the versions according to Tab. 8.1. The light gray bandwidths are the 90 / 95 % confidence intervals of 37 years according to Chap. 7.1. All results correspond to Freiburg (1961 – 1980).

nealing can only be applied if the evaluation of the objective function is fast. Therefore, the complete PDF is only calculated for the initial time series and after each swap just the incremental changes on the PDF are computed and evaluated. The simple spell definition reduces the possible influences of the swaps on the distribution to eight cases. The changed value can either be wet or dry and each swap can occur at four different positions with respect to the spell. The position of the changed value can be at the beginning or at the end of a spell, two spells can be merged or one spell can be split. All eight cases have their specific influences on the wet and dry spell distributions. That is, the change of the PDF after each swap can be reduced to the identification of the cases of both swapped values, followed by the respective changes of the PDF.

I perform an optimization using version $p4x$ including the wet and dry spell distribution as target parameters. However, the additional optimization on dry spells does not lead to further improvements, and thus I discuss only the results of the optimization of the wet spells in the following.

In Chap. 3.1 I showed a significant difference of the wet spell distribution related to different measurement devices. Therefore, I use two different target distributions in the optimization. The target distribution of version $p4xf$ is calculated using time period I (1961 – 1980), whereas in version $p4xf2$ the distribution of period III (2007 – 2012) is applied.

Fig. 8.9 (left panel) shows the reduction of the wet spell frequency of version $p4xf$ and $p4xf2$. Not only the frequency of small spells is significantly reduced, but also their respective precipitation sum can be improved (Fig. 8.9 left panel). This demonstrates that in the NiedSim algorithm not too many parameters need to be part of the objective function. If the most important ones are found, these constraints can lead to a realistic structure while other parameters are fitted indirectly. Unfortunately, spells with a long duration (> 18 h) can not

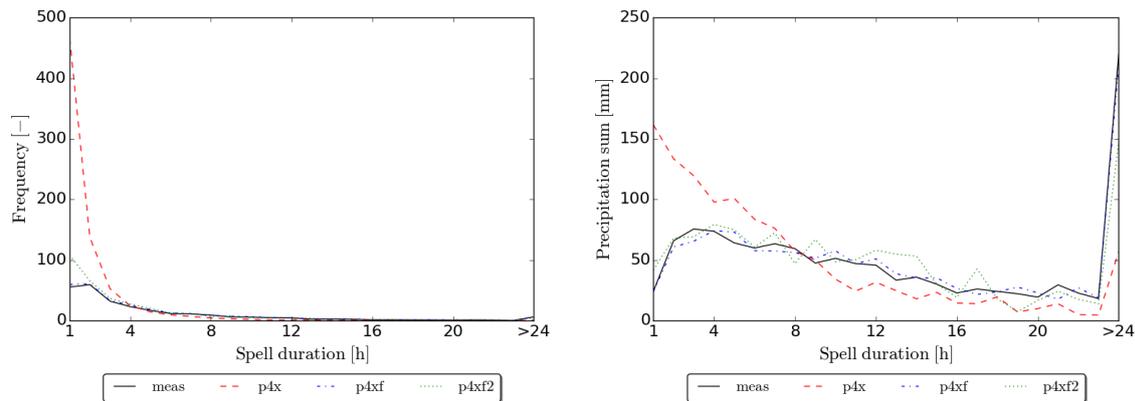


Figure 8.11.: Histogram of wet spell durations (left panel), where the spells are separated by at least one dry hour as well as the precipitation sum contributed by each duration (right panel). The four different cases are measured data and the versions according to Tab. 8.1. All data refer to Kempten (1961 – 1980).

be improved very well and still exhibit an underestimated precipitation sum.

If the influence with respect to the CSOs is regarded (Fig. 8.10), the optimization of the wet distribution can only slightly improve the time series of Freiburg. As expected, the impact is largest for CSO durations of fast responding systems with sensitive structures (11, 31, 51 showing small specific discharges q_s). As the optimization of the wet spell distribution focuses rather on medium and small values, its effect on the partial series is negligible and, thus, not explicitly shown.

Interestingly, the statistics of the smooth time series of time period I ($p4xf$), which are supposed to be less realistic compared to period III ($p4xf2$), can lead to better results in the NiedSim optimization. Although the effect on CSOs is only small (Fig. 8.10), it shows that more realistic measurements may not necessarily improve the optimization of NiedSim. Nevertheless, the small improvements do not justify the usage of unrealistic data to tweak the optimization. Instead, one should use realistic input data and try to improve the algorithm considering additional statistics.

The relevance of the wet distribution for Freiburg is less important, whereas for other locations, the impact can be very high. Figure 8.11 shows the wet spell distribution of the station Kempten. Not only the number of short spells are overestimated more in version $p4x$ compared to Freiburg, but also their respective precipitation sums are far too high. The underestimation of the precipitation sum of spells with longer durations is also more emphasized. These issues lead to serious underestimations of the CSO statistics in Kempten (Fig. 8.12). In contrast to Freiburg, not only the CSO duration, but also the CSO volume is affected. Like in Freiburg, the optimization of the wet spell frequencies is also able to improve the distribution of the precipitation sums indirectly (Fig. 8.11 right panel).

Another side aspect should be mentioned. For Freiburg (and Weiden, although not explicitly shown), spells with a long duration (> 18 h) that contribute a considerable precipitation

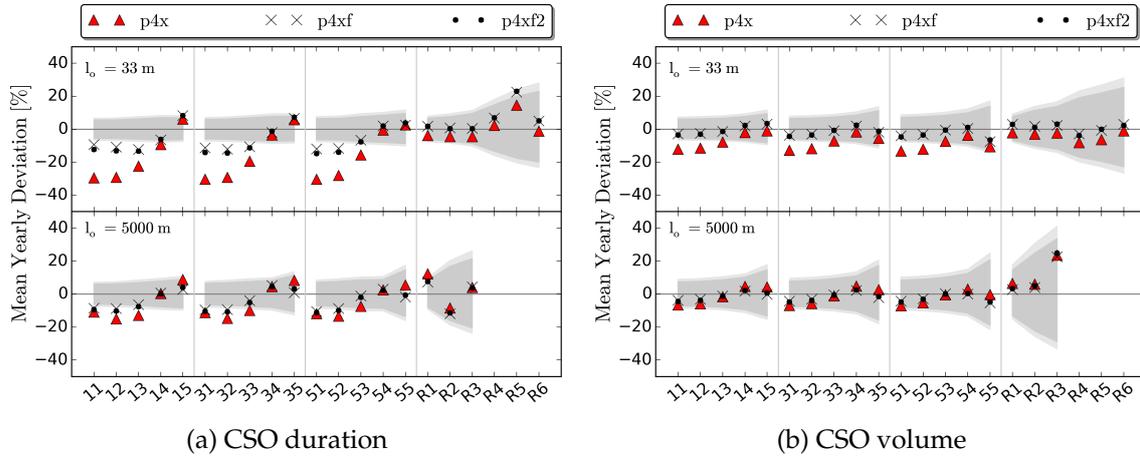


Figure 8.12.: Difference in CSO statistics of measured time series and hourly optimized time series using NiedSim (mean of 10 runs). The four different cases are measured data and the versions according to Tab. 8.1. The light gray bandwidths are the 90 / 95 % confidence intervals of 37 years according to Chap. 7.1. All results correspond to Kempten (1961 – 1980).

sum can not be optimized indirectly if the frequency distribution is fitted (Fig. 8.9 right panel). Therefore, I test other optimizations, in which the distribution of the precipitation sums is explicitly optimized. Neither an individual optimization, nor an optimization in combination with the frequencies, nor applying higher weights on longer durations could significantly increase the precipitation sum of longer spells. For Kempten, however, a good optimization of the precipitation sum can already be achieved by an optimization on the frequency distribution. A major difference between the locations are the annual precipitation sums (Kempten: ≈ 1200 mm/year, Freiburg ≈ 900 mm/year, Weiden ≈ 700 mm/year). This suggests that a minimum volume of precipitation must be available such that the algorithm is able to optimize certain characteristic of the time series. For Freiburg (and Weiden), the precipitation sum of spells larger 18 h might be too small to be properly distributed within long duration classes. Although the underestimation of the long spells seems not to be responsible for the remaining issues of the CSO duration (Kempten shows similar systematic deviations, although the precipitation sums could be fitted), it demonstrates that the same algorithm may work differently for different locations. I will discuss the impact of this result further in the outlook (Chap. 11).

In summary, an optimization of the classified wet spell frequency is able to improve the respective precipitation sums of each class indirectly and, thus, leads to a more realistic temporal structure. Depending on the regarded location, the effect of the improved structure on applications is relevant and leads to smaller deviations in the CSO statistics. The less realistic digitized paper chart records of period I have a positive impact on the synthetic modeling due to their influence on the hourly wet spell distribution. Although I would not recommend to use the more unrealistic data, they, if used, at least do not worsen the precipitation simulation.

8.3. Final discussion on the hourly optimization

I introduced the Poisson process in the setup of the initial time series of NiedSim as well as the wet spell frequency as additional parameter in the objective function. I validated the synthetic time series of both modifications indirectly using the partial series as well as the artificial combined sewer system. The time series could be improved with respect to large and extreme values, but still tend to underestimate the partial series of the measured data slightly. This can be explained with the fact that no explicit optimization on extreme values were applied. However, the deviations were rather small (less than 10 %) for most parts of the partial series.

The indirect validation using the sewer also showed a clear improvement of the NiedSim optimization algorithm. Nevertheless, systematic underestimations in the CSO duration for certain structures could still be observed. Therefore, two questions remain: how realistic are the time series and how should the residual deviations be handled? For an answer, I regard different event-based statistics in the following as they are not used as input statistics for NiedSim.

During the further investigations I will analyze the data of Kempten because the deviation in the CSO duration is more pronounced compared to Freiburg. I will use the three versions developed in the previous chapters that built directly on each other: version *a*, mainly based on the autocorrelation; *p4x*, including the Poisson process; and *p4xf2*, additionally optimized on the wet spell distribution.

At first, univariate event statistics are investigated. For this purpose, I take the definition of the wet spells used before as event criteria, that is, I simply separate spells by one dry hour $ts_e = 1$ h (Fig. 8.13a). As I already illustrated in the last section, if the intermittency is not considered (*a*, *p4x*), a very large number of spells are generated that are too small and too short. The introduction of the wet spell distribution (*p4xf2*) is not only able to improve the frequency, but indirectly all other spell statistics.

A simple spell definition using $ts_e = 1$ h can be easily implemented in the NiedSim optimization and might be important for very fast responding systems. In most sewers, however, the concentration time is larger than 1 h, such that the discharge time series is smoother compared to the initial rainfall series due to retention and attenuation processes. Therefore, many details in the wet spell rainfall statistics do not change the discharge series in the sewer and, thus, are irrelevant for the CSO behavior. Furthermore, small individual values as well as little events cannot affect applications because they do not contribute to run-off at all.

Consequently, I apply a second event criterion that is more related to sewer applications. The criterion uses the minimum time with zero precipitation that separates two events (ts_e) in addition with the minimum intensity (i_e) and the minimum event depth (d_e). That is, 5 min values that are smaller than i_e are treated as zeros and events smaller than d_e are not counted as events. The parameters I use here are $ts_e = 8$ h, $i_e = 0.1$ mm and $d_e = 1.5$ mm. The selection of the thresholds (i_e and d_e) must be chosen carefully to avoid a cutting of too much data. I select the criterion $i_e = 0.1$ mm as it is the resolution of a conventional

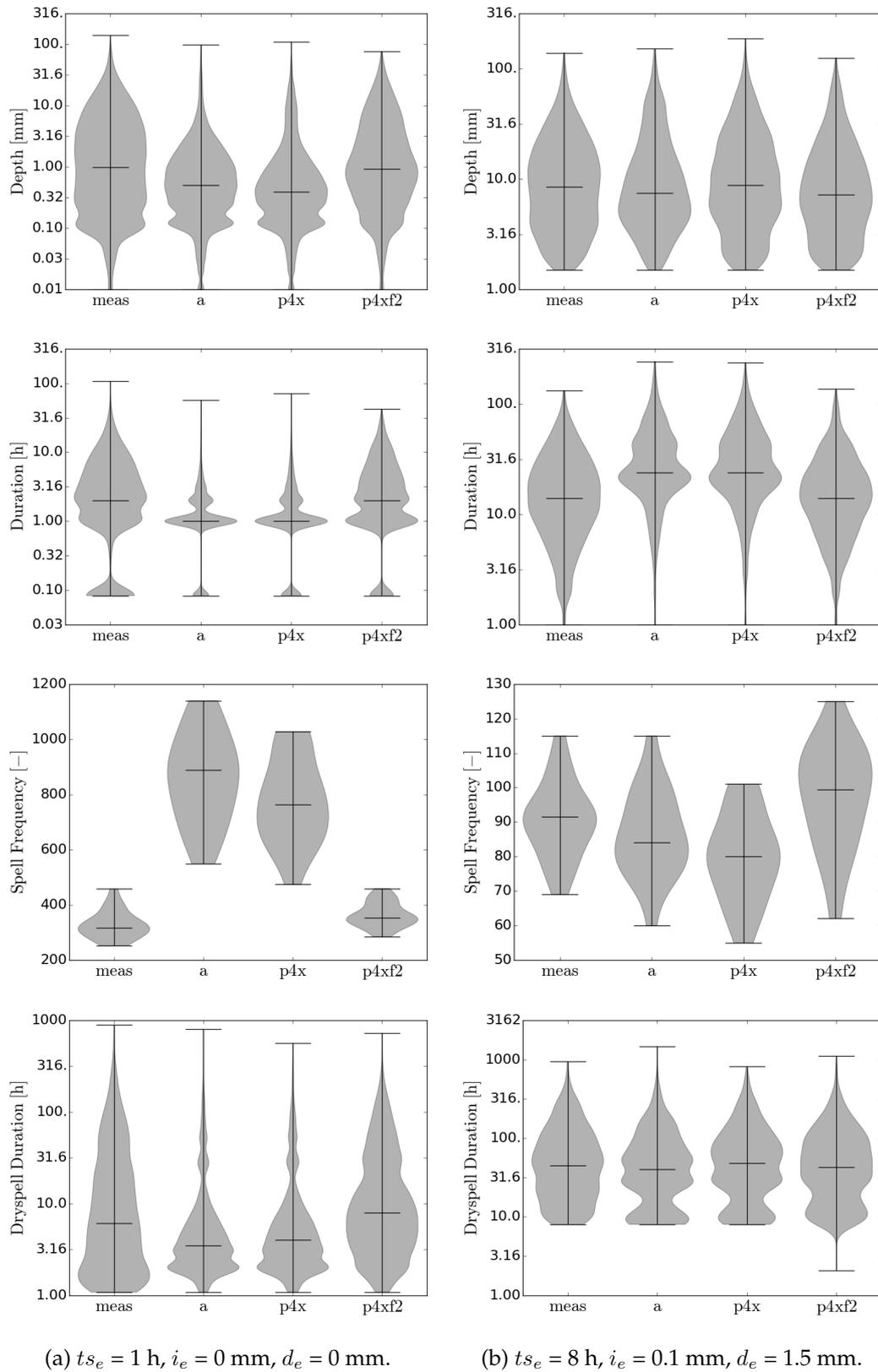


Figure 8.13.: Distribution of mean annual spell statistics including min., max. and median. The versions are defined according to Tab. 8.1. All data correspond to Kempton (1961 – 1980).

tipping-bucket; I define $d_e = 1.5$ mm such that it is in the range of the depression storage of the run-off generation (see Chap. 5.2.1).

For event-based sewer simulations t_{s_e} is usually chosen considering the response time of the system, which is the flow time in the catchment and in the sewer, and can be in the range of several hours. The reason is to assure the independence of events if not the entire time series is simulated. However, if the event statistics are used for validation purposes or to infer about the subsequent CSO behavior as in my case, the selection of t_{s_e} becomes more difficult. It is always a trade-off between independence of events and a detailed representation of the internal intensity variability as explained in Chap. 2.2.2. That is, a short t_{s_e} reflects the characteristics of short spells of high intensities best, which are relevant for fast responding sewers, whereas a long t_{s_e} is necessary to characterize the long term relationship of rainfall spells that are important for highly buffered systems. As I used a very short $t_{s_e} (= 1$ h) as the first criterion before, I apply a rather long $t_{s_e} = 8$ h as second criterion now.

Figure 8.13b illustrates that the differences in the spell statistics of the three versions are less pronounced for $t_{s_e} = 8$ h compared to $t_{s_e} = 1$ h (Fig. 8.13a). Nevertheless, a systematic improvement can still be observed for version *p4xf2*. That is, the optimization of the distribution of the wet spell frequencies can indeed improve large parts of the temporal structure. Nonetheless, even the best version (*p4xf2*) shows not identical distributions in the spell statistics compared to the measured data. However, none of the spell statistics seem to be noticeable enough to explain the underestimation of the CSO duration in Fig. 8.10 and Fig. 8.12.

A problem with the univariate spell statistics applied so far is that they cannot reflect the interdependence structure between the statistics. For example, it can not be judged whether long precipitation spells contribute large rainfall sums or if most precipitation is related to short spells. Both examples can show the same univariate spell distribution.

A possibility to explain the remaining deviations in the CSOs could be the consideration of a bivariate classification. For this purpose, I first classify the precipitation spells according to spell depth and duration and calculate the frequency for each class (corresponding to a two dimensional histogram). Finally, I compute the mean annual deviation of 10 synthetic time series from the measured time series and plot them in Fig. 8.14 (left panel). I take the mean of 10 series to obtain more robust characteristics because the frequencies of some classes can be rather small. Additionally, I analyze the deviation of the cumulative precipitation sum of all spells falling in the respective class (similar to the investigation of the univariate case) and illustrate them in Fig. 8.14 (right panel). In Fig. 8.14 I use the spell separation criteria $t_{s_e} = 1$ h, $i_e = 0$ mm and $d_e = 0$ mm, whereas Fig. 8.15 illustrates the results for events with a separation criteria $t_{s_e} = 8$ h, $i_e = 0,1$ mm and $d_e = 1,5$ mm.

The plots are directly related to the ones of the univariate case. Figure 8.11 shows the marginal distribution of the spell durations, which can be obtained from the absolute data in Fig. 8.14 by a summation along the y-axis.

Several conclusions can be drawn from the classification. In the frequency case (Fig. 8.14 left panel), the largest deviations occur in the lower left corner, that is, for small and short spells. As already mentioned, due to the skewness of precipitation, large deviations in the

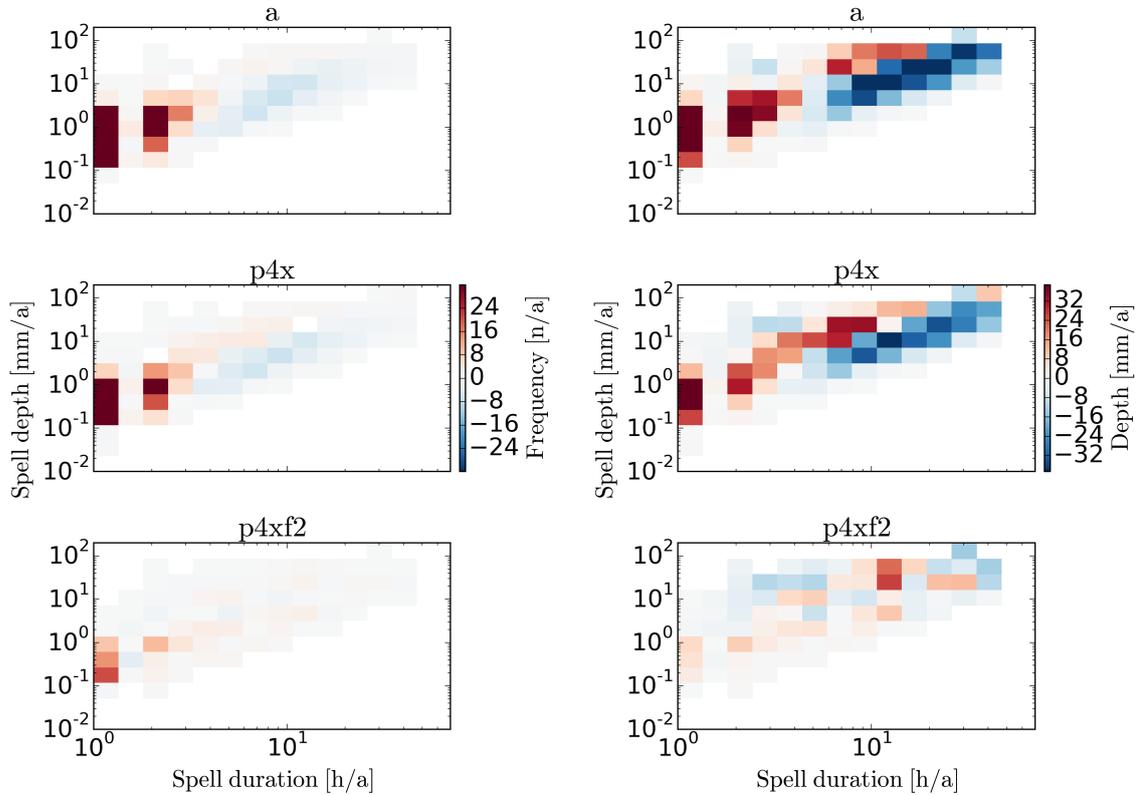


Figure 8.14.: Dependence structures of precipitation spells separated with $ts_e = 1$ h, $i_e = 0$ mm, $d_e = 0$ mm. The spells are classified in two dimensions according to depth and duration with logarithmic classes. Each class shows the deviation of 10 simulated time series (see Tab. 8.1) from the measured series with respect to the mean annual frequency (left panels) and the mean annual accumulated precipitation depth (right panels).

frequency of short and small events are less relevant as their total spell sum is small. On the other hand, a few missing long and heavy rainfall events are hardly noticeable in the frequencies, but can lead to serious underestimations of the total sum (Fig. 8.14 right panel). If one optimizes on the univariate wet spell distribution $p4xf2$, not only the frequency, but also the cumulative sums are improved. This result is not new and has already been shown in the univariate case. However, Fig. 8.14 shows, that the optimization also improves the interdependence structure of spell duration and depth of the bivariate case. Furthermore, even if the second spell separation criteria is applied, where the spells are separated by $ts_e = 8$ h (Fig. 8.15), the optimization of the wet spell distribution shows a clear improvement.

Nevertheless, even a bivariate event analysis with different event separation criteria does not indicate pronounced errors that can be directly related to the underestimation of the CSO duration. There is a tendency, that very large and long spells (upper right corner) are underestimated, but the assumption cannot be clearly verified for Freiburg (A.6) or Weiden

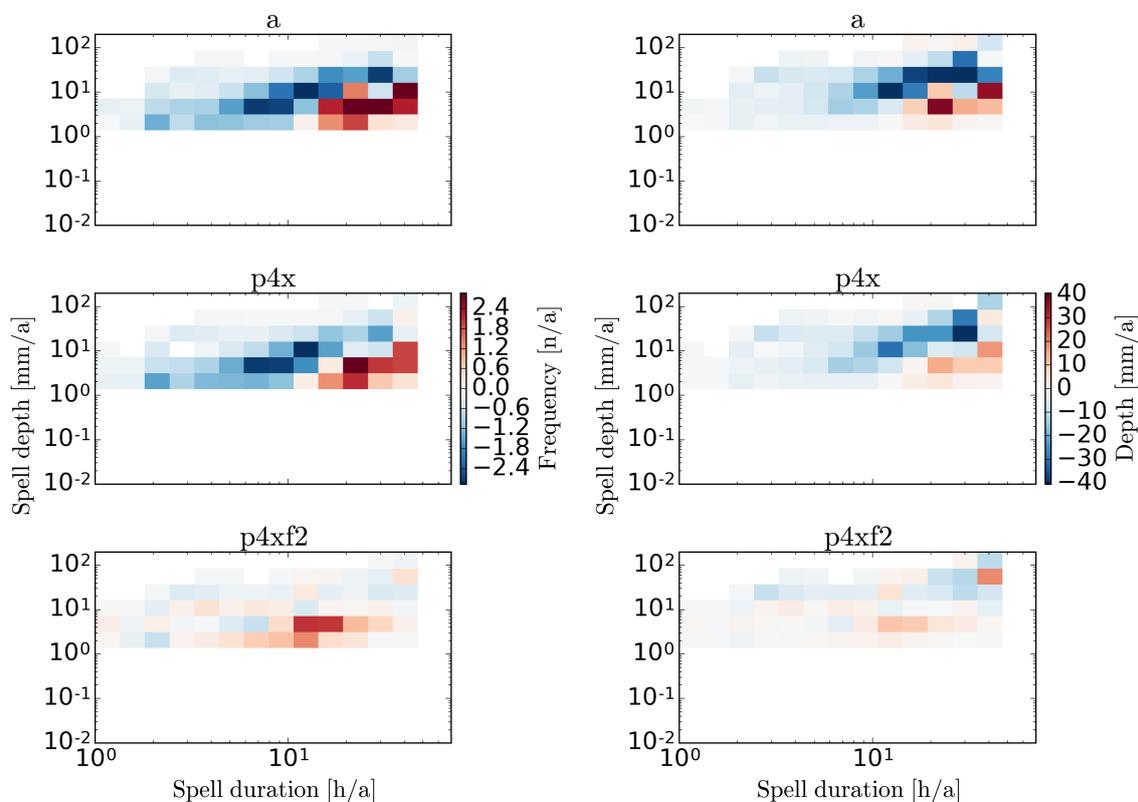


Figure 8.15.: Dependence structures of precipitation spells separated with $ts_e = 8$ h, $i_e = 0.1$ mm, $d_e = 1.5$ mm. For a description see Fig. 8.14.

(not shown as similar to Freiburg).

One faces several general issues if one tries to infer from event-based statistics on sewer applications. First, the event characteristics change depending on the regarded temporal aggregation and, even more critically, on the event definition. For example, $ts_e = 1$ h shows long, but small events being underestimated for version *a* and *p4x* (Fig. 8.14), whereas in $ts_e = 8$ h, these spells are overestimated (Fig. 8.15). Additionally, the actual influence on the sewer of an event depends on its multivariate characteristics (frequency, depth and duration). Furthermore, the precipitation is highly skewed, such that a few events can influence the results of the application significantly. Finally, the relevance of different rainfall characteristics depends on the regarded overflow structure and on the definition of the CSO event. At the end, using event-based statistics one deals with various possible combinations that all show the same result, but from a different perspective. It might help to get a general idea of the complexity of the rainfall-runoff-overflow processes, but it is not easy to keep track on the really important characteristics.

Finally, one could ask, why I did not optimize on the multivariate distribution in NiedSim. One crucial point is the skewness of the rainfall events that can lead to small mean annual spell frequencies of a class of less than once per year. However, NiedSim is optimized on

an annual base which makes it difficult to reflect such return periods. Another challenge is the fast implementation in the swapping routine. An optimization on the univariate volume instead of the frequencies increases the computational time already by a factor of two.

In conclusion, the improved hourly optimization by using a Poisson process and the wet spell durations is able to capture most rainfall statistics including the partial series. It shows only small deviations in the CSO volume for a large bandwidth of overflow structures used in practice. For the CSO duration, issues can be observed depending on the location and the regarded overflow structure. However, as the CSO volume is the most important parameter for the combined sewer simulation, the optimization of the hourly temporal structure can be supposed to be applicable for practice.

9. Analysis of the disaggregation of NiedSim

The final step of NiedSim is the disaggregation of the hourly time series into a 5 min time series. In this section, I investigate the performance of the disaggregation algorithm. In order to avoid an influence of the hourly optimization, I use measured 5 min time series as initial time series of the disaggregation. The intra-hourly temporal structure of the measured series is destroyed by using uniformly distributed 5 min values within each hour (block rainfall). The temporal structure is then reconstructed with the disaggregation algorithm of NiedSim. By comparing the disaggregated and the original time series in 5 min resolution, one can judge the quality of the disaggregation. Any deviation between the series can be related purely to the disaggregation algorithm because their hourly temporal structure is identical.

Like in the hourly optimization, I use the partial series and the artificial sewer for an indirect validation. As the intra-hourly temporal structure is not relevant for long overland flow paths due to the smoothing effect (see Fig. 7.6 left panel), I use only the sewer with $l_o = 33$ m in this section.

9.1. Motivation

In the analysis of precipitation for different time periods, I illustrated a large influence of the measurement device on many statistics (Chap. 3.3). In particular, the impact on the 5 min statistics can be very large. As they serve as targets in the optimization of NiedSim, the different time periods also affect the performance of the disaggregation. Figure 9.1 illustrates large deviations in the CSO statistics of structures without storage if parameters of different time periods are considered in the objective function. As these structures reflect large and extreme values, the influence can also be observed in the IDF-curves as shown later.

One could argue that the parameter set for the disaggregation of one time period might not fit to the characteristics of hourly rainfall of another time period, and thus a temporal cross validation using different time periods will always yield biased results. However, the influences of the annual temporal structure is rather low as shown by the q-q-transformation before (Fig. 7.4), whereas the input statistics exhibit significant deviations. As they can be referred to different measurement devices (Chap. 3.3), it is very likely that different measurement instruments are the reasons for the deviations.

NiedSim2.x was developed around the year 2000. Consequently, the input statistics were dominated by digitized paper records (period I). However, by using an increasing amount of data from newer devices (period II and III), the data set has become more and more

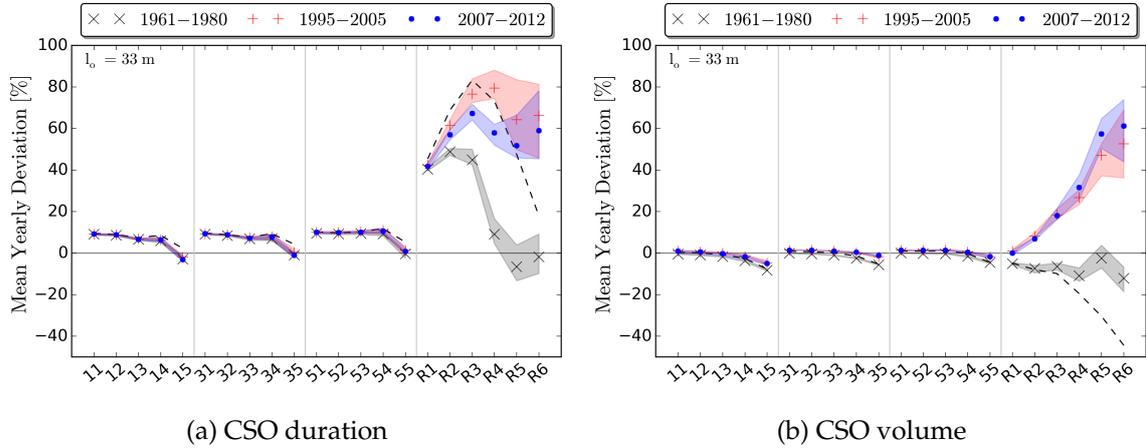


Figure 9.1.: Influence of the parameters of different time periods on the mean annual CSO statistics of Freiburg. The plot illustrates the deviation of three NiedSim disaggregated time series from the original, measured time series from 1961 – 1980. The disaggregated series were obtained from the hourly aggregated, measured series, which were disaggregated using target parameter sets of different time periods. The bandwidths represent the minimum and maximum of 10 different disaggregation runs. The dashed line shows the results of the random intensity structure serving as benchmark (Chap. 7.2.1). The calculation is based on the disaggregation version *w1s2* defined in Chap. 9.2.1.

inhomogeneous, and led to inconsistencies between the disaggregation algorithm and the new data sets.

As explained in Chap. 3, the time period III represents the natural precipitation characteristics the best. Unfortunately, only short periods of period III are available for many stations such that the influence of sampling errors can be high. On the other hand, the longer time period I shows a less realistic temporal structure, and thus the largest deviations in the target statistics. As more and more old measurement devices are replaced by new ones reflecting the characteristics of time period III, they will sooner or later dominate the local statistics. Therefore, the goal of my investigation is the development of a robust disaggregation algorithm based on the measured statistics of period III, where the deviations shown in Fig. 9.1 are reduced.

As for the stations Kempten, Freiburg and Weiden the data of time period III is very short, I use another location, Augsburg, in the following where period III covers sixteen years (1997 – 2012).

9.2. Modifications of the disaggregation algorithm

In NiedSim2.x the autocorrelation, scaling parameters and 5 min exceedance probabilities are considered as well as the 5 min and hourly maxima obtained from KOSTRA (see

Chap. 4.4). In order to be able to track the influence of each parameter, I chose the initial parameter set to be as small as possible. In a preliminary work, *Gunzenhauser* (2016) determined the autocorrelation as well as scaling parameters to be crucial for the disaggregation, and that the 5 min exceedance probabilities are less important. Although the 5 min maxima can be relevant, I do not use them here because they are based on an extreme value analysis. As explained before, the target of my study is obtaining a robust mean annual temporal structure without considering the extremes explicitly. Consequently, I start my investigation using the disaggregation with autocorrelation and scaling parameters as targets in the objective function.

The target scaling parameters of the disaggregation in NiedSim2.x are calculated with the entire range of aggregations (5 min – 24 h). Due to the scaling breaks illustrated in Chap. 3.4.1, the scaling parameters for the entire range are smaller compared to the ones calculated using only intra-hourly aggregations (Fig. 3.4b, marked dots). A decreased scaling parameter of the third moment b_3 is related to relatively less uniform values at the 5 min aggregation compared to the hourly aggregation. Consequently, the intra-hourly precipitation values are optimized such that they are more heterogeneous compared to the observed ones. The result is a large overestimation of the partial series in NiedSim2.x. In order to avoid this issue here, I calculate the scaling parameters used as targets in my study only based on intra-hourly aggregations.

In NiedSim2.x the objective function includes the autocorrelation only at a 5 min aggregation for lag 1 – 10 and the 5 min weighted autocorrelation for lag 1 with a higher weight on winter. Here, I consider the weighted autocorrelation with a higher weight on summer as well as additional aggregations {5, 10, 15, 30, 45, 60 min} of lag 1 and 2. The changes are supported by the following facts that are based on *Gunzenhauser* (2016) as well as on additional preliminary studies by myself. First of all, the focus of the optimization should be on summer because the high resolution data particularly affects combined overflow structures without storage. These CSOs are caused by short, but heavy precipitation events mainly occurring in summer. Additionally, I could not find an advantage from optimizing both, the autocorrelation and weighted autocorrelation. Consequently, I use only the weighted autocorrelation as target parameter. Furthermore, the influence of the autocorrelation of different temporal aggregations and/or lags was investigated to be rather small. However, the correlations at higher aggregations seem to be slightly better optimized if more aggregations are considered rather than an increased number of lags.

9.2.1. Autocorrelation and scaling parameters

In order to investigate the influence of the scaling parameters and the weighted autocorrelation on the disaggregation, I use four different optimization functions. In the first two versions, both parameters are analyzed individually, that is, version w and s include only the weighted autocorrelation and scaling parameters respectively. As I will show later, both parameters cannot be optimized simultaneously, and thus the optimization depends on the weighting factors ω of the parameter (Eq. 4.1). The lower ω , the higher the actual weight due to the standardization of the individual parts of the objective function. In version $w1s2$

and *w2s1* both parameter sets are optimized at the same time, but with different weighting factors. For example, in the version *w1s2* the weighted autocorrelation is optimized with $\omega = 1$ and the scaling parameters with $\omega = 2$, and hence the version reflects a focus on the correlation.

In Tab. 9.1 the statistics of the resulting time series of four disaggregation versions are shown in addition to the measured data (*meas*) as well as a randomly changed series (*rand*), where the objective function does not include any parameter. The version *rand* is the benchmark for the disaggregation as explained in 7.2.1.

The optimization of the weighted autocorrelation generally works (version: *w*) because the deviations of the autocorrelation are much lower compared to the random version. However, differences of around $\pm 10\%$ for the selected correlations show that the optimization is not able to reflect the input statistics perfectly even if no other target parameters are considered. Additionally, one can notice an improvement of the scaling parameter of the second and the third moment although they are not directly optimized. This indicates that both parameters are not independent of each other.

The same characteristics can be noticed, if only the scaling parameters are optimized (version: *s*). They also can be improved compared to the random optimization, but still show some considerable deviations from the target scaling parameters. Again, a dependence with the weighted autocorrelation is observed. However, the effect is negative such that the autocorrelation, in particular, at high resolutions ($r_1^{(5min)}$) is even worse compared to the random case.

An explanation can be given with the characteristics of the scaling parameters as interpreted in Chap. 3.4. To restate, the smaller the scaling parameter, the more heterogeneous the small aggregation (5 min) is compared to the large aggregation (1 h). Consequently, homogeneous block rainfall yields the highest possible scaling parameters. The scaling parameters must be reduced during the disaggregation to increase the heterogeneity of the intensities for smaller aggregations. A low scaling parameter can be obtained if only a few values within the respective hours are highly increased. However, if only a few values are enlarged without any other condition like in version *s*, these values will be scattered, and thus the autocorrelation will be underestimated.

For version *w* the optimization behaves differently. The measured autocorrelation at high temporal resolutions is considerably high (here, $r_1^{(5min)} = 0.7$), which is lower compared to block rainfall, but higher compared to the random optimization ($r_1^{(5min)} = 0.49$). Therefore, an optimization of the weighted autocorrelation (version *w*) tries to balance the autocorrelation, which is easier to obtain if the intensities are not very heterogeneous. Consequently, the resulting scaling parameters are higher compared to measured data.

The strong dependence of correlation and scaling parameters is confirmed if both are included as target parameters in the optimization. If the weight is on the weighted autocorrelation (*w1s2*) or on the scaling parameters (*w2s1*), the resulting deviations of the correlations are larger for both combinations compared to the version where only the weighted autocorrelation is optimized (*w*). For the resulting scaling characteristics, the picture is different.

Version	abs [-]						rel [%]				
	meas	rand	w	s	w1s2	w2s1	rand	w	s	w1s2	w2s1
$r_1^{(5min)}$	0.70	0.49	0.63	0.27	0.60	0.51	-30	-10	-61	-14	-27
$r_1^{(15min)}$	0.51	0.69	0.48	0.41	0.42	0.36	37	-5	-20	-16	-28
$r_1^{(30min)}$	0.41	0.65	0.46	0.50	0.38	0.34	57	12	21	-6	-17
b_1^{summer}	0.70	0.86	0.86	0.77	0.85	0.80	22	22	9	21	13
b_1^{winter}	0.70	0.82	0.82	0.74	0.81	0.77	18	18	6	17	11
b_2^{summer}	1.31	1.63	1.54	1.26	1.45	1.32	25	17	-4	11	1
b_2^{winter}	1.46	1.57	1.57	1.45	1.51	1.46	8	7	-1	3	0
b_3^{summer}	1.73	2.35	2.04	1.45	1.79	1.69	36	18	-16	4	-2
b_3^{winter}	2.04	2.21	2.22	1.88	2.02	2.04	9	9	-8	-1	0

Table 9.1.: Overview of weighted autocorrelation r and scaling parameters b of measured time series (*meas*) and disaggregated time series using different versions of NiedSim. The versions use different target parameters in the objective function: *w*: only weighted autocorrelation; *s*: only scaling parameters; *w1s2*: both parameters with a focus on weighted autocorrelation; *w2s1*: both parameters with a focus on scaling. The version *rand* reflects a random change without considering any target parameters. All data refer to Augsburg (1997 – 2012).

As soon as the autocorrelation is additionally included, the scaling parameters of some moments can even be optimized better compared to the version using scaling parameters only (*s*) depending on the chosen weighting factors ω .

To summarize the section, it is already challenging to optimize the autocorrelation or the scaling parameters individually. Furthermore, in contrast to the hourly optimization, both parameter sets cannot be optimized independently of each other in the disaggregation and, thus, the weighting factor ω is relevant.

9.2.1.1. Validation of scaling parameters and autocorrelation

In the last section, I illustrated the ability of the NiedSim optimization to reproduce the statistics of scaling parameters and weighted autocorrelations. In the following, I analyze their indirect influence on combined sewer systems as well as on the partial series, on the probability of zero precipitation p_0 and finally, on event-based statistics.

CSO statistics The performance of the four different disaggregation versions with respect to combined sewer systems is illustrated in Fig. 9.2, where the mean annual deviation in CSO statistics from the original time series are plotted. Large deviations can be observed for structures without storage (R1 – R6), whereas the influence is rather low for structures with storage (11 – 55). I have already shown these characteristics in Chap. 7.2.1.

Figure 9.2b indicates that version *w1s2* performs the best with respect to CSOs. The deviations in the volume are reduced compared to the random intensity structure (the benchmark

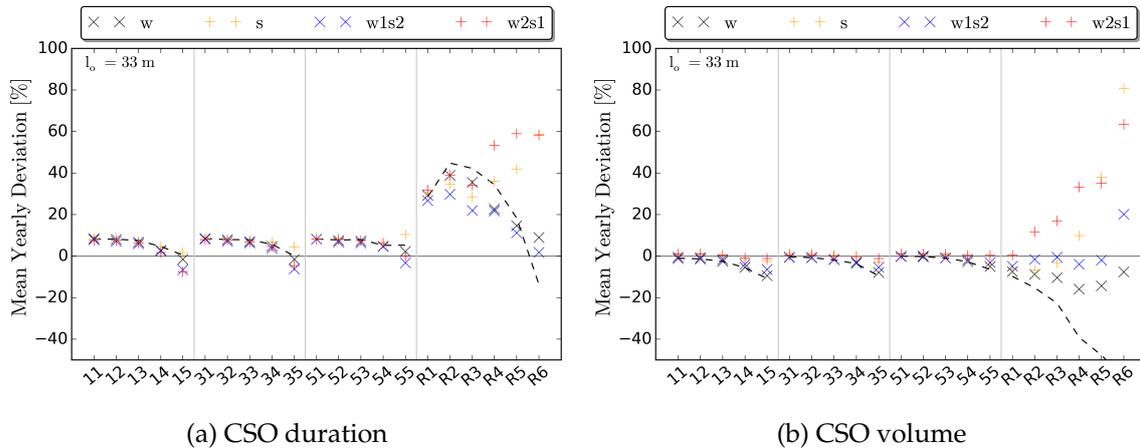


Figure 9.2.: Mean annual deviations of CSO statistics of different versions of disaggregated time series from the original time series. The same versions as in Tab. 9.1 are shown here. The dashed line shows the results of a random intensity structure (*rand*) serving as benchmark explained in Chap. 7.2.1. All data refer to Augsburg (1997 – 2012).

for the disaggregation), in particular, for structures that are sensitive to extremes due to high specific discharges (R2 – R6). For other structures as well as for the CSO duration only minor improvements can be observed.

If the focus in the optimization is on the scaling parameters (*w2s1* and *s*), the volume for structures without storage is highly overestimated. The reason for the large CSO volume is the overestimation of large and extreme values as I will explain in the analysis of the partial series in the following.

Partial series In order to analyze the influence of the disaggregation on large and extreme values, I investigate the partial series. Figure 9.3 clearly indicates a large influence of different target parameters in the objective function.

If only the weighted autocorrelation (*w*) is optimized, the extreme values are in general underestimated. That is, the values are structured in a rather homogeneous way such that no large values occur because nothing forces them to peak. On the other hand, if only scaling parameters are considered (*s*), the five minute values are highly overestimated. As explained before, the lower the scaling parameters the more heterogeneous the structure is, and the more large values occur. In the optimization, the scaling parameters must be reduced compared to the initial time series (block rainfall). The reduction can be achieved very simply with a large increase of a few single 5 min values. The result is an overestimation of the 5 min extremes while higher durations are still underestimated.

If both parameters are optimized with a focus on scaling (*w2s1*), the extreme values (above the red line) are fit rather well except for 5 min durations. However, if one considers additionally large, but not extreme values (below the red line), a high overestimation is visible.

An explanation could be that the optimization of the scaling parameters increase some 5 min values while at the same time the weighted autocorrelation forces the neighboring values to increase as well, which leads to an overestimation at higher durations. On the other hand, if the weight ω is on the weighted autocorrelation ($w1s2$), the tendency of the correlation to homogenize the time series seems to prevent an enlarging of too many neighboring 5 min values.

In all cases, difficulties are visible in the representation of the entire range of the partial series. As shown in Chap. 3.5, the partial series for very short durations show high variabilities including possible influences from measurement devices. As the true underlying population is unknown, it is difficult to judge the optimization algorithm with such uncertain references.

The characteristics of the extended partial series (below the red line) are an indicator of the behavior of the CSO volume of structures without overflow in Fig. 9.2b. An underestimation leads to a reduced CSO volume, whereas an overestimation causes too large overflow volumes compared to measured data. However, no relationship between CSO duration and partial series can be noticed. This again shows that the CSO duration cannot fully be described by the partial series. Therefore, other measures are used in the next section to investigate the temporal structure of smaller values not reflected by the IDF-curves.

I would like to mention an aspect of scaling parameters here, which is their vulnerability to outliers. Eq. 2.2 indicates that the higher the moment is, the more it is influenced by extreme values. Therefore, *Lombardo et al.* (2014) recommends to use only the first two moments. Unfortunately, the NiedSim algorithm is not able to produce good results with this configuration. Instead, the partial series is highly overestimated if only the first two moments are part of the objective function (plots not explicitly shown).

Probability of zero precipitation The most simple way to describe precipitation is the occurrence of rainfall. It can be regarded as a Bernoulli experiment with two possible results: *rain* or *no rain*. In literature, the occurrence of zero precipitation p_0 is often analyzed. Due to its obvious relationship to the duration of rainfall spells, such an investigation seems to be appropriate here to get some further insights into the errors of the structure of disaggregated time series. Unfortunately, p_0 is not a very robust parameter because of the boolean characteristic of the outcomes. The impact of small precipitation values on p_0 is high, and its uncertainty is further increased due to large measurement errors for small precipitation intensities (*Lanza et al.*, 2005). On the other hand, even substantial deviations in p_0 of two time series need not necessarily lead to large impacts on subsequent applications. For example, the occurrence of many small, individual values in simulated time series (synthetic drizzle) highly affect p_0 although their influence on sewer simulation is rather small. Therefore, I investigate the probability of values exceeding a certain threshold p_e instead of analyzing p_0 . Such thresholds must be chosen very carefully because the influence of small measured values on CSO statistics in high resolution time series can be significant (see Chap. 7.2.3).

In my study I use a threshold of 0.1 mm, that is, $p_e(H^{(j)} > 0.1)$ to avoid issues with drizzling. Additionally, I analyze only hours with considerable rainfall ($H^{(h)} > 1.0$ mm).

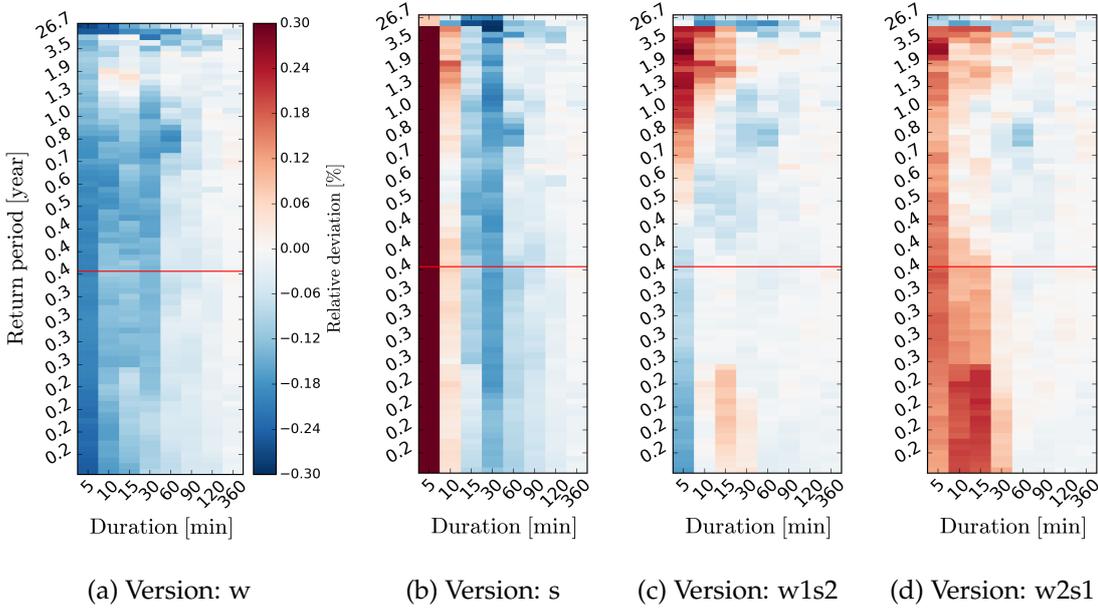


Figure 9.3.: Relative deviation of partial series of different versions of disaggregated time series from the original time series. Return periods above the red line are usually considered for extreme value analysis. All data and versions refer to the ones defined in Tab. 9.1.

Version	abs [-]						diff [-]				
	meas	rand	w	s	w1s2	w2s1	rand	w	s	w1s2	w2s1
$p_e^{(5min)}$	0.562	0.507	0.502	0.520	0.499	0.472	-0.06	-0.06	-0.04	-0.06	-0.09
$p_e^{(15min)}$	0.843	0.944	0.945	0.932	0.942	0.923	.10	.10	.09	.10	.08
$p_e^{(30min)}$	0.924	0.995	0.997	0.995	0.997	0.997	.07	.07	.07	.07	.07

Table 9.2.: Probability of zero precipitation $p_e(H^{(j)} > 0.1)$ calculated only for hours with at least 1.0 mm rainfall. All data and versions refer to the ones defined in Tab. 9.1.

In both cases, the disregarded values are likely to have a low impact on CSOs as shown in Chap. 7.2.3.

Table 9.2 gives an overview of p_e of the different versions used before. At the 5 min aggregation the probability of exceedance is slightly underestimated in all simulations, whereas at higher aggregations p_e is overestimated. The reduced $p_e^{(5min)}$ indicates that the hourly precipitation is spread to fewer 5 min values compared to the original data. However, these values are very scattered within the hour because at a higher aggregation the number of exceeding values increases. In other words, the 5 min values are not structured with a clear defined beginning and end, but spread across the entire hour. The logical consequence is an increase of the CSO duration as illustrated in Fig. 9.2.

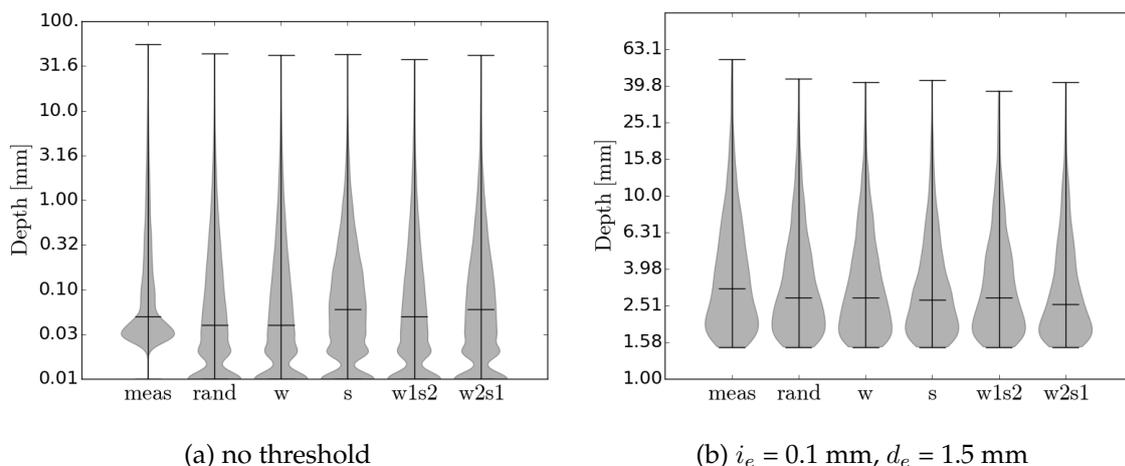


Figure 9.4.: Distribution of event statistics including minimum, maximum and median. In both cases ts_e is set to 5 min. All data and versions refer to the ones defined in Tab. 9.1.

Event statistics Having these results in mind, I investigate the more complex event-based statistics. I apply the same event criteria as introduced for the hourly optimization, that is, the minimum time with zero precipitation that separates two spells (ts_e), the minimum intensity (i_e), and the minimum event depth (d_e). However, as the event statistics should allow to infer on the intra-hourly structure, a $ts_e < 1$ h must be selected.

Due to the short ts_e , the relevance to apply thresholds (i_e and d_e) is even more important. If spells are separated with $ts_e = 5$ min, many very small values occur that are too small to influence discharge characteristics, but can highly affect event statistics. Fig. 9.4 is an example of the distribution of the event volume calculated with and without thresholds. The difference between the event definition criteria seems to be tremendous as the entire distribution changes. However, most of the deviations are artifacts caused by different resolutions (0.03 for most observed measurements in Augsburg; 0.01 for the simulation). Consequently, one needs to apply some intensity thresholds if one would like to infer conclusions from high resolution event-based statistics.

The definition of ts_e is also crucial for the calculation and interpretation of event characteristics as already mentioned several times in my study. In the following, I apply two criteria, $ts_e = 5$ min (a) to reflect very short intense spells, and $ts_e = 30$ min (b) to focus on the interconnection of several values. Figure 9.5 illustrates the distributions of event statistics obtained with both definitions. In case (a), deviations in the distribution between the versions can be noticed, but it is difficult to relate them to the CSO behavior. An interim time of $ts_e = 5$ min is too short to relate CSO characteristics with event statistics.

In case (b), the distributions of the synthetic time series are all very similar to each other. Furthermore, the systematic overestimation in the event duration of all versions can explain the increased CSO duration compared to the measured series in Fig. 9.2. Almost no precipitation event occurs that is shorter than 60 min. Therefore, the conclusion of the previous section can be confirmed: the disaggregation algorithm is not able to generate individual,

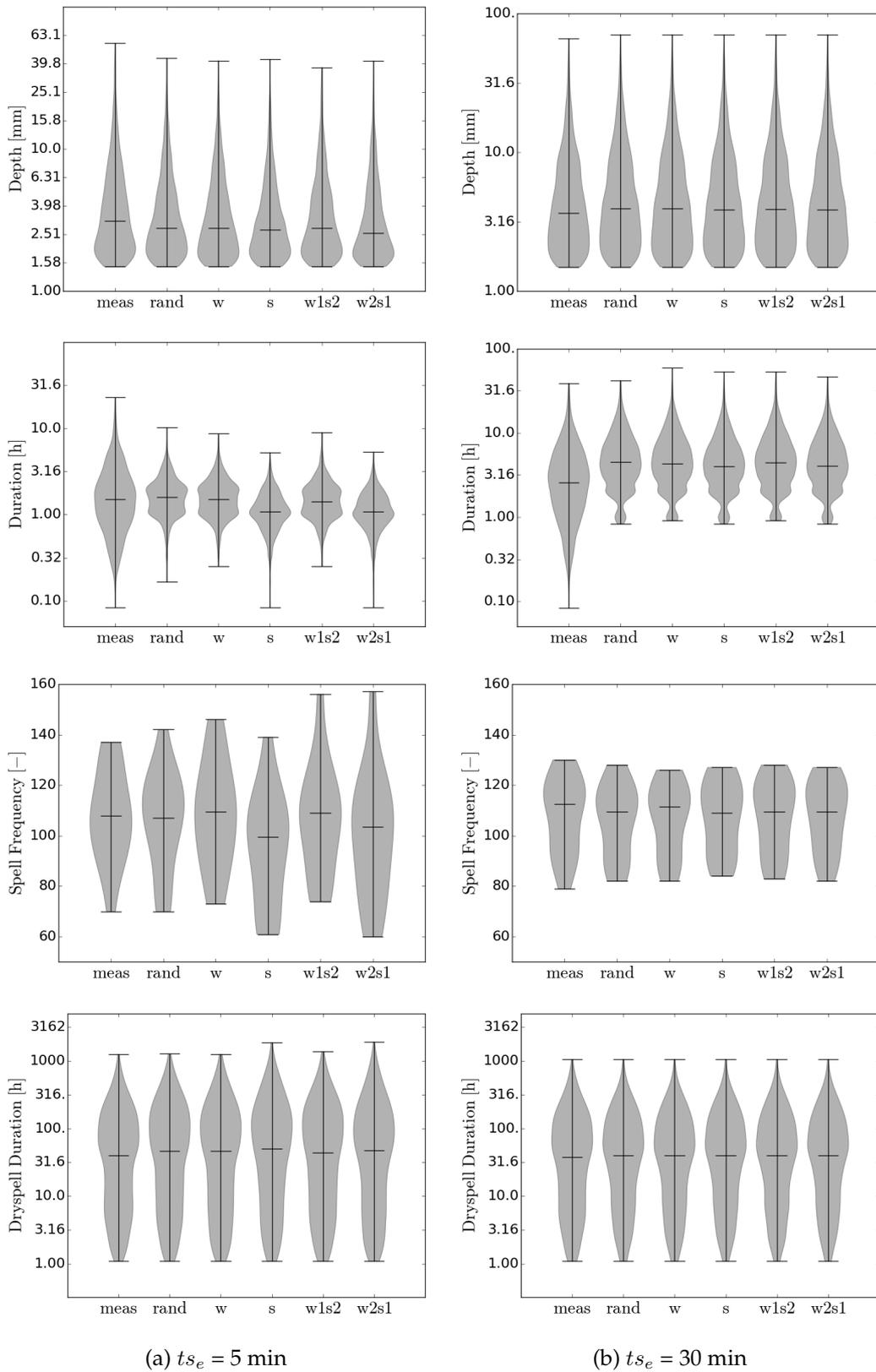


Figure 9.5.: Distribution of mean annual event statistics including min., max. and median. Both cases are set up with $i_e = 0.1 \text{ mm}$ and $d_e = 1.5 \text{ mm}$. All data and versions refer to the ones defined in Tab. 9.1.

short spells within an hour.

Conclusion The disaggregation shows deficiencies if scaling parameters and weighted autocorrelations are used as target parameters. As both parameters influence each other, they cannot be optimized simultaneously. The scaling parameter tries to optimize highly heterogeneous structures, whereas the autocorrelation rather smooths the structure. The result is that the algorithm does not optimize short, realistic spells with a high intensity and well defined start and end points. Instead, a random background noise of small values is generated that is interrupted by some very large 5 min values scattered within the hour without forming proper events. These characteristics lead to a systematic bias in the CSO duration. Therefore, I introduce an additional parameter, the entropy, in the objective function in the following.

9.2.2. Entropy

A possibility to describe the 5 min distribution of each hour is the entropy H_S as shown in Chap. 2.2. A high entropy indicates that the hourly sum is uniformly distributed in the 5 min values, whereas a low value ($H_S = 0$) shows that the total sum is concentrated on just one single 5 min value.

In the following, I compare the entropy distribution of the measured and simulated time series. For this purpose, I calculate the hourly entropy using 5 min values according to Eq. 2.5. However, instead of regarding each hour individually, I compute the entropy of two consecutive hours. The idea is to enable NiedSim to optimize the entropy across two hours if it is part of the objective function (which will be the case later on). This way, issues with spells separated due to the hourly discretization might be reduced. Therefore, H_S can take values from the interval $[0, \log_2(2 \cdot 12) = 4.58]$. Like in the analysis of p_0 , I analyze only hours with at least 1.5 mm of precipitation to avoid unimportant hours to cover relevant characteristics.

Figure 9.6a illustrates the histogram of entropies of measured data, of the randomly changed time series and of a disaggregated time series (*w1s2*). Because of the skewed hourly entropy and for optimization reasons, I classify the data not equidistantly with 24 classes, but such that the frequencies of the measured series are approximately uniformly distributed. Obviously, the randomly disaggregated time series (*rand*) shows large differences in entropy compared to the measured one. Most of the hours have entropies between approximately three and four, whereas classes with very high or very low values do not occur. This is not surprising because from an optimization point of view, both ends indicate extreme cases. It is neither likely that the hourly precipitation amount structures itself uniformly across all 5 min increments, nor combined in just one 5 min value. Consequently, most 5 min values are optimized in a balance between both extreme cases, such that the hourly entropy histogram peaks somewhere between. Two peaks occur because I analyze the entropy of two consecutive hours, but each hour is disaggregated individually. The first peak, appears if only one hour has a precipitation sum larger than zero, whereas the second peak reflects cases where both hours are wet.

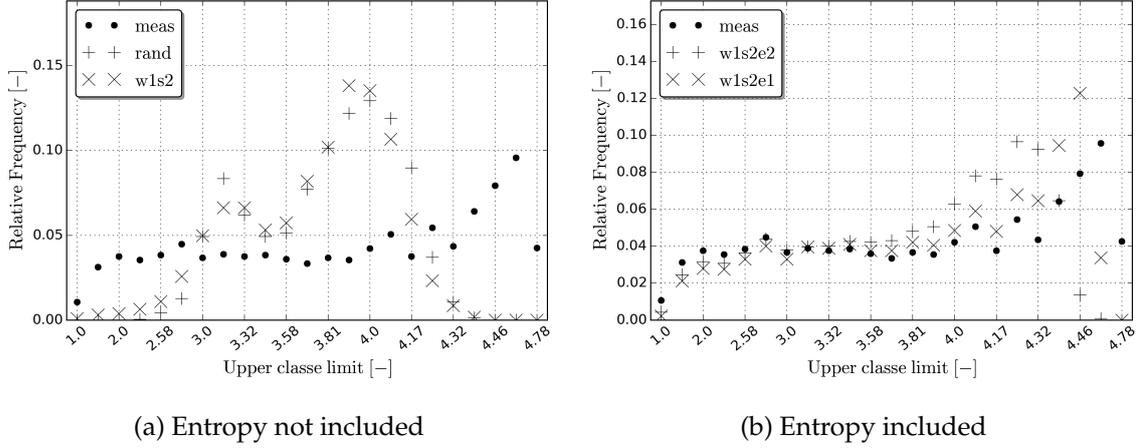


Figure 9.6.: Scatter plot of entropies and hourly precipitation depths of different time series. All data and versions refer to the ones defined in Tab. 9.1. Version $w1s2e1$ and $w1s2e2$ are equal to $w1s2$, but with the entropy as additional target parameter.

If the time series is disaggregated only with weighted autocorrelation and scaling parameters in the objective function ($w1s2$), the distribution of the entropy is almost the same compared to the random case. That is, neither the scaling parameters, nor the autocorrelation can influence the hourly entropy significantly. If the relative frequencies of the entropy are included in the objective function of the disaggregation (version $w1s2e1$ and $w1s2e2$), a similar entropy distribution compared to the measured data can be obtained except for very large entropies (Fig. 9.6b). Both versions differ only in the optimization weight ω of the entropy, where $\omega = 1$ in version $w1s2e1$ and $\omega = 2$ in $w1s2e2$.

However, if the CSO statistics are analyzed (Fig. 9.7), there is almost no difference between $w1s2$, $w1s2e1$ and $w1s2e2$. There are two reasons, why improved entropies do not directly affect the CSO results as explained in the following.

The first reason can be illustrated if one regards a scatter plot of the hourly entropy and its respective precipitation sum. From top to bottom in Fig. 9.8, the number of targets in the optimization function increases. In the case of a random disaggregation (a), the hourly entropies gather between three and four for the reasons I have already explained. An optimization of the autocorrelation (b) only affects hours with a large precipitation sum (> 8 mm), for which the entropy is slightly reduced ($H_S < 3$). However, most of the hours are not affected at all, which is not surprising because the autocorrelation is dominated by large values as I illustrated before. An additional optimization of the scaling parameters (c) yields a similar result compared to (b), but the large hours tend to be even more peaked ($H_S < 2$) and fit better to the measured structure. Nevertheless, the scaling parameters influence particularly large hours, whereas the bulk is similar to the random optimization.

Figures (d) and (e) illustrate the case where both parameters are optimized in addition to the entropy. The results highly depend on the weighing factor of the optimization. If a large focus on the entropy is applied (Fig. 9.8 (d), version $w1s2e1$), the time series is strongly, but unrealistically, forced towards the measured entropy distribution. One can clearly observe

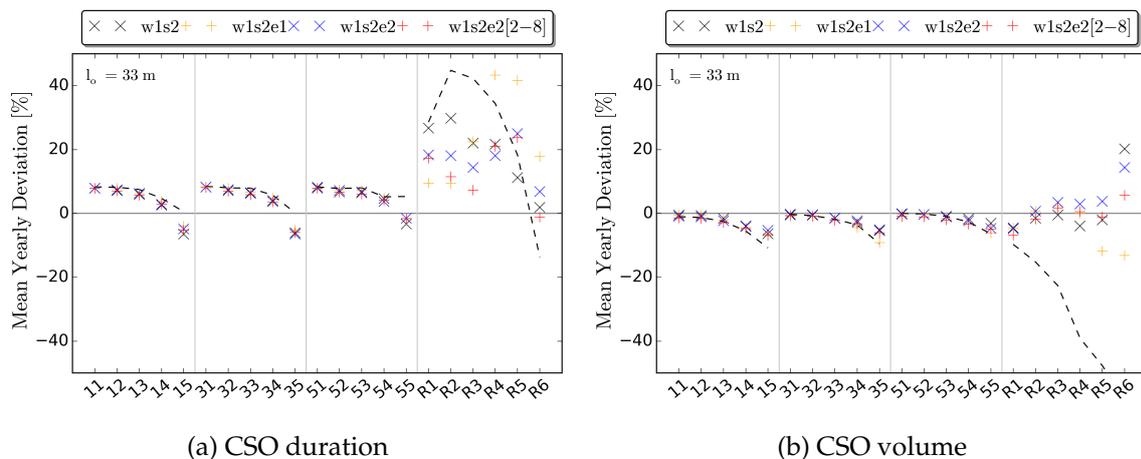


Figure 9.7.: Mean annual deviations of CSO statistics of different versions of disaggregated time series from the original time series. The same precipitation data was used for the sewer simulation as shown and explained in Fig. 9.8.

that the entropies are not homogeneously distributed within each class, but cluster at one boarder – the boarder, which is closer to the likelier class for random optimization (here $3 < H_S < 4$). Furthermore, only hours with a rather small sum (< 8 mm) are optimized towards more extreme entropy structures ($H_S < 3$ or $H_S > 4$). This is a logical result, as the higher the precipitation sum, the more single rainfall increments must be changed in the optimization, to achieve an entropy structure which is significantly different from being random. For smaller sums, more extreme structures can be obtained more easily and, thus, are more likely after the optimization. From a application point of view, version *w1s2e1* is problematic because large hours are relatively homogeneously distributed. However, heterogeneous hours with a rather small entropy are more critical for sewer systems.

If the focus on the entropy is reduced (Fig. 9.8 (e), version *w1s2e2*), the structure of large hours is dominated by autocorrelation and scaling parameters such that hours with less homogeneous structures can occur. Here, the optimization of the entropy forces hours with a very low intensity to reflect low entropies.

One could try to force the optimization of the entropy only for hours with medium intensities. Therefore, I consider hours with 2 – 8 mm rainfall in the optimization of the entropy (Fig. 9.8 (f), version *w1s2e2[2-8]*). The results highlight a strong clustering at the class borders, that is, towards the likeliest entropy classes (3 – 4) and towards smallest hourly sums (2 mm). In conclusion, it is difficult to optimize the structure to reflect very high or very low entropy values. Even if the time series is optimized with a focus on entropy, the result is an unrealistic entropy distribution and the distribution is only forced towards the target entropies artificially.

The second issue, why the CSO duration can not be improved by the optimization of the entropy, is its fitting at just one aggregation. The entropy describes if rainfall at a sub-scale is homogeneously distributed or if single peaks occur. However, it cannot describe the temporal order of the values. In other words, from a low hourly entropy one can conclude that

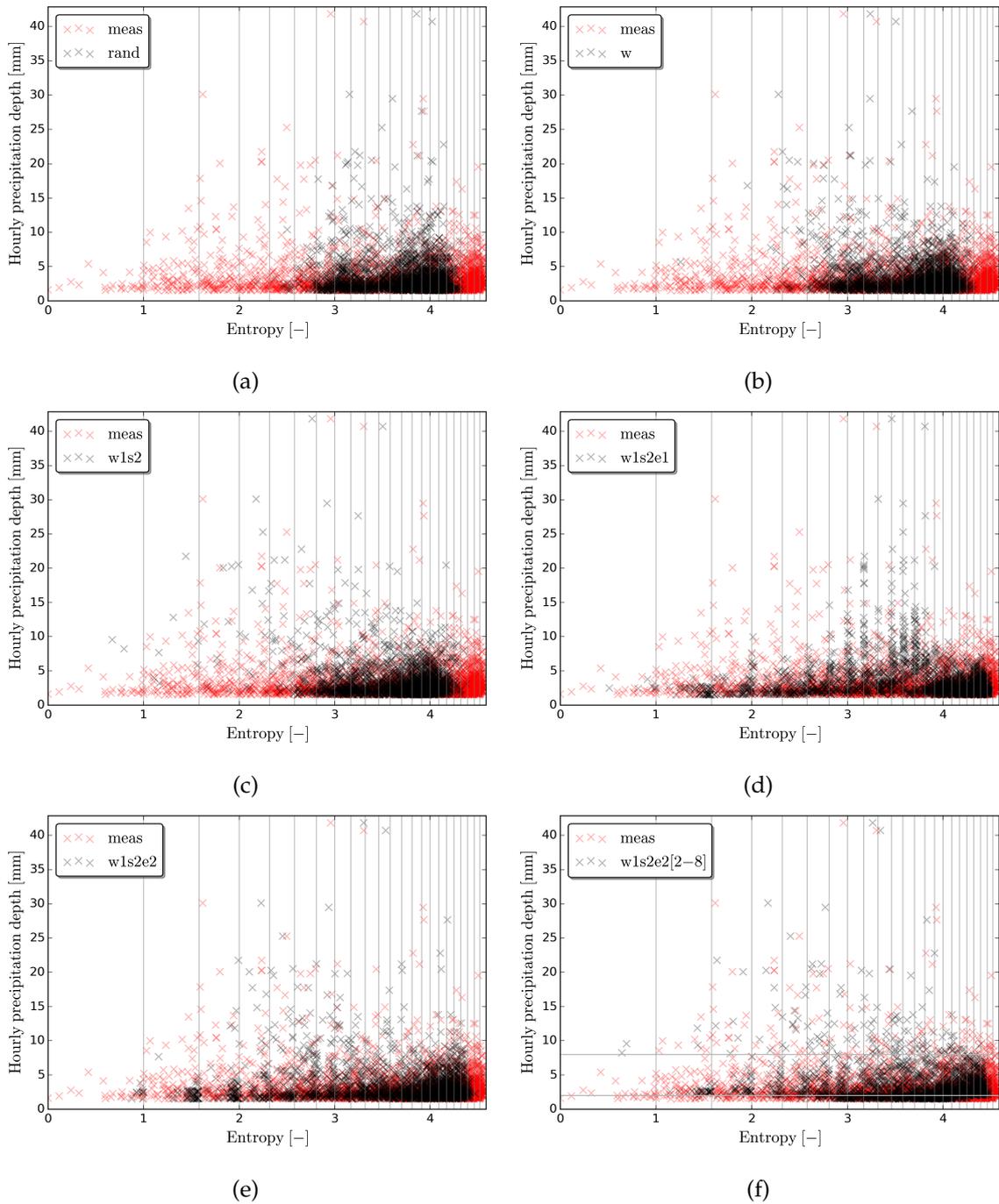


Figure 9.8.: Scatter plot of entropies and precipitation sum of different time series. All data and versions refer to the ones defined in Tab. 9.1. In Fig. (d, e), the versions are equal to $w1s2$, but with the entropy as additional target parameter and different optimization weights. Version $w1s2e2[2-8]$ is the same as in (e), but the entropy is only optimized for hourly depths between 2 – 8 mm.

5 min values are heterogeneously distributed, but not if they form a short, connected event, or if they are scattered across the hour. Only if the entropy of several sub-hourly time scales were considered, such an information could be reflected. As Fig. A.7 illustrates, the optimization of the entropy at the 5 min aggregation does not improve the entropy structure at larger aggregations. In particular, low entropies are missing and indicate a too homogeneous structure. One could try to optimize on the entropy at additional aggregations. However, an improvement of the time series seems to be doubtful because the optimization of the entropy at the 5 min aggregation is already difficult. Additionally, the quality of the NiedSim optimization is usually worse, the more dependent variables are included. This would be the case if several time scales were regarded. Therefore, I do not consider this task in this study.

9.3. Final discussion on the disaggregation

A realistic intra-hourly structure is only important for applications where high intensities are crucial. A wrong intensity distribution can lead to systematic errors, in particular, for design purposes as well as for CSO structures without storage of fast responding sewers. Consequently, the IDF-curves and the indirect validation of the fast sewer ($l_o = 33$ m) are the relevant validation tools for high resolution synthetic time series.

Unfortunately, measured partial series are highly uncertain at high resolutions and, additionally, might be influenced by measurement devices. Therefore, it is difficult to use empirical partial series for the validation of extremes of synthetic series at high resolutions. Therefore, I did not focus on the partial series in the validation of the disaggregation. Furthermore, I illustrated in Chap. 7.2.2 that the influence of extreme values on CSO statistics at high resolutions is negligible such that they could be adjusted in a post-processing step without significantly changing the overflow characteristics.

I demonstrated with the indirect validation that the high resolution temporal structure has a rather low impact on combined sewer structures with storage (deviations < 10 %) even for fast reacting sewers. Consequently, only for structures without storage, and only for fast reacting sewers, the disaggregation is important.

An optimization of the 5 min temporal structure with NiedSim using the weighted autocorrelation and the scaling parameters cannot account for a realistic intra-hourly temporal structure. The NiedSim disaggregation generates a few high 5 min intensities, which are even larger than the observed ones. The algorithm is not able to reproduce distinct precipitation spells with large and medium intensities. In particular, the NiedSim disaggregation lacks in the reconstruction of events with well-defined starts and endings.

An additional optimization on the entropy could not further improve the time series with respect to CSO statistics. Furthermore, an alternative optimization on the distribution functions of different aggregations led to worse results with an unrealistic temporal structure as well as missing large and extreme intensities (not explicitly shown before).

Nevertheless, the disaggregation is not useless because the disaggregated time series show better results compared to ones that are randomly changed.

Finally, I suggest to use the parameter set summarized in Tab. 9.3 for the generation and disaggregation in a new version of NiedSim.

Parameter	Calculation period	Aggregation [min]
1 h generation		
ϕ^* : hourly distribution	period I – III	60
ϕ : monthly sum	monthly, period I – III	-
ϕ : daily exceedance probability of 0, 1, 5 mm/day	annual, period I – III	1440
ψ^* : Poisson process	period I – III	60
ψ : autocorrelation (lag 1)	annual, period I – III	60, 120, 180, 360, 720, 1440
ψ : weighted autocorrel. (lag 1)	annual, period I – III	60, 120, 180, 360, 720, 1440
ψ : wet spell frequencies	period III	60
5 min disaggregation		
ψ : weighted autocorrel. (lag 1, 2)	period III	5, 10, 15, 30, 45, 60
ψ : scaling exponent (1 st – 3 rd moment)	winter / summer, period III	5, 10, 20, 30, 60

Table 9.3.: Final parameter set used in NiedSim3 for the 1 h generation and 5 min disaggregation. ϕ^* and ψ^* are only used for the setup of the initial time series and not for the optimization.

10. Additional properties of rainfall relevant for sewer systems

In Chap. 2, I introduced statistics that almost all share one common property, that is, the temporal reversibility. Most of the proposed statistics would not change if the underlying precipitation time series was reversed in time. However, the temporal direction of a precipitation time series matters for some applications like sewer simulations as I will show in the following sections.

Irreversibility of precipitation time series is not new and has been observed for decades. For example, the previously mentioned Huff curves show that the cumulative distributions of the first and the fourth quarter are different (*Huff, 1967*). The temporal direction dependence is also considered in some applications like in synthetic design storms based on IDF-curves used for the design of sewer networks (*DWA-A118, 2006*). Nevertheless, to my knowledge, there is only one statistical measure available in literature that describes and quantifies the temporal irreversibility (*Yu et al., 2013*). This measure is based on precipitation events. Other approaches only detect whether a time series is reversible or not (*Giannakis and Tsatsanis, 1994; Steuber et al., 2012*), but cannot quantify its degree.

In the following, I will introduce a new measure, the asymmetry of dependence, that detects, describes and quantifies the irreversibility of continuous precipitation time series. It is based on the work of *Sugimoto et al. (2016)* who applied the theory for discharge time series. I add several modifications compared to them which are necessary for the usage in rainfall time series. Before I go into details, I would like to remark that the entire Chap. 10 is based on *Müller et al. (2017a)*.

10.1. Theory of asymmetry of dependence

The asymmetry of dependence characterizes irreversibility and is based on the theory of copulas. Therefore, I will shortly review both terms before I introduce the actual measure.

Irreversibility A precipitation time series $X(t)$ is reversible according to *Lawrance (1991)* if

$$X(t), X(t+1), \dots, X(t+k) \stackrel{d}{=} X(t+k), X(t+k-1), \dots, X(t) \quad (10.1)$$

for any lag $k = 1, 2, \dots$ and time step $t = \pm 1, \pm 2, \dots$, where $\stackrel{d}{=}$ denotes identical in distribution. Consequently, the characteristic of reversible time series is related to the joint distributions of time series separated by lag k . For many natural processes the equation does not hold, and thus the resulting time series cannot be reversed in time and is irreversible.

Irreversibility can be easily illustrated using a plotted time series. If the increasing and decreasing graphs show different slopes, the time series is irreversible. Such a characteristic is well known from hydrographs where the discharge increases rapidly, followed by a long lasting decrease after the peak (the recession curve). The same behavior, but less pronounced, can be detected in hyetographs (Fig. 10.1). For illustration purposes, I show an event where the visual asymmetry is very pronounced. In most other precipitation events the temporal order is less distinct, but can still be detected with the asymmetry of dependence as shown later.

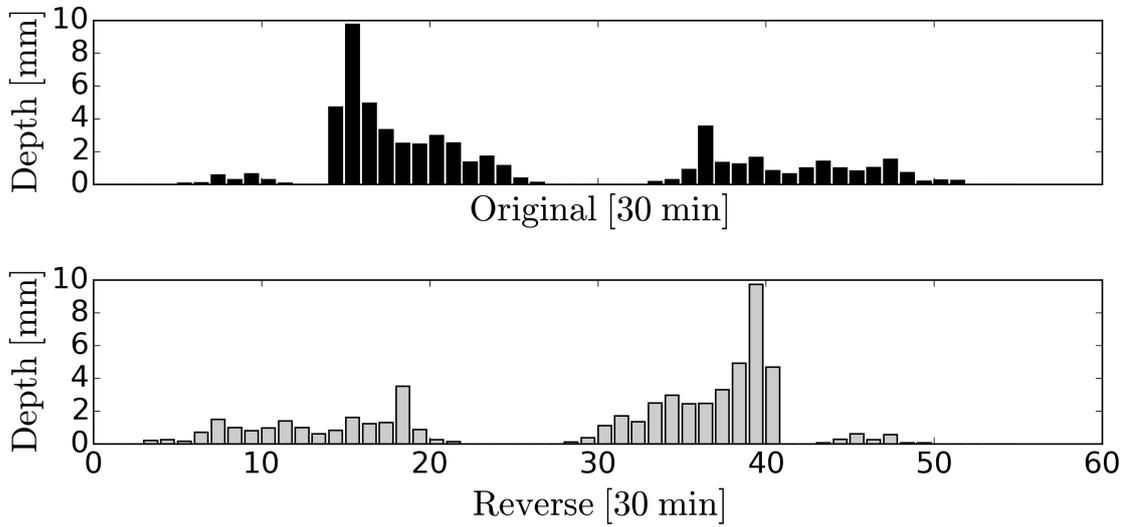


Figure 10.1.: Original precipitation time series (length: 30 hours) and its reversed form in a temporal resolution of 30 minutes in Bavaria (figure from Müller *et al.*, 2017a).

Copulas The asymmetry measure as I introduce here is based on a bivariate copula. A bivariate copula is a joint cumulative distribution function, where both variables are uniformly distributed on the interval $[0, 1]$.

$$C(u_1, u_2) = P(U_1 \leq u_1, U_2 \leq u_2) \quad (10.2)$$

The copula describes the pure dependence structure of both variables in the rank-space and is not influenced by their marginal distributions. Consequently, statistics based on copulas are robust to outliers. As the reversibility is a property of the joint distribution function, the bivariate auto-copula $C(u_1, u_2)$ with $U_1 = F(X(t))$ and $U_2 = F(X(t+k))$ yields information about the reversibility, where k is the time lag. In case of a reversible time series

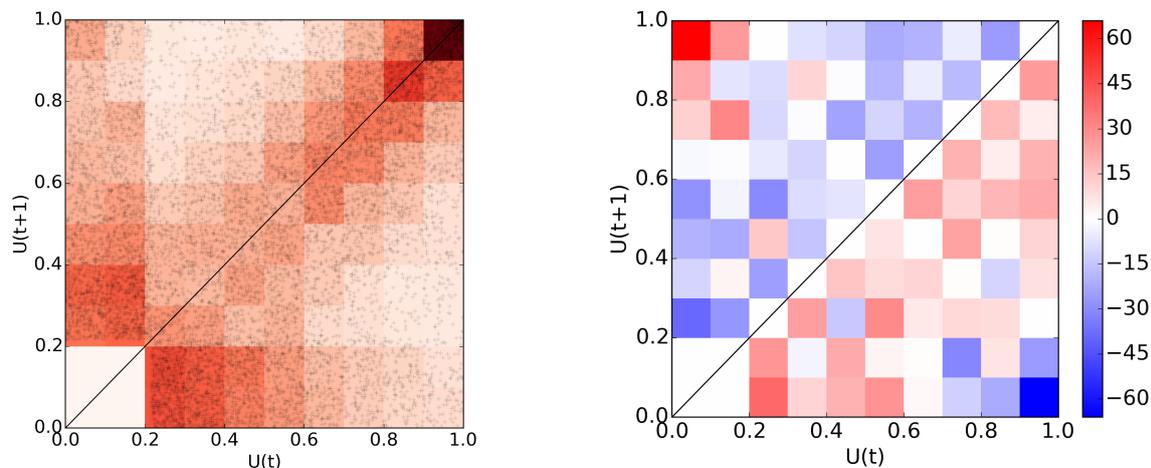


Figure 10.2.: Left: Empirical copula density of a precipitation time series. Each cross indicates a pair $\{U(t); U(t + 1)\}$. The colored squares indicate the number of pairs falling in the respective square. The darker the color, the higher the frequency. Right: Differences between upper-left and lower-right triangle, $c(u_1, u_2) - c(u_2, u_1)$. Data used in both figures: Kempten, May – Sept., 1999 – 2012, resolution: 30 min, threshold: $x_t = 0.1 \text{ mm}$ (figure from Müller *et al.*, 2017a).

$X(t)$, the resulting copula is symmetric with respect to the bisecting line. Otherwise the time series is reversible and the copula asymmetric with the copula density

$$c(u_1, u_2) \neq c(u_2, u_1) \quad (10.3)$$

Figure 10.2 (left panel) is an example of an empirical auto-copula density of a precipitation time series with $k = 1$. An empirical copula can be generated by plotting the ranks $U(t)$ and $U(t + 1)$ in a scatter plot. The density of the copula is reflected by the colored squares indicating the number of pairs within each square. The density is highest close to the bisecting line and illustrates the relatively high correlation of precipitation values at an aggregation of 30 min. If values at a time step are high (low), the values at the next time step are most likely high (low respectively) as well. An important deviation can be noticed in the upper left corner where the density is rather high. The pairs falling in this square show low values at the current time step t , but high values at the next time step $t + 1$. If one reconsiders Fig. 10.1, these values represent a fast increase of precipitation intensity and can be related to a sudden start of a heavy rainfall event (e.g. convective cell). The reversed case (lower right corner), which is an abrupt stop of an intense spell, occurs much less frequently.

The asymmetric characteristic can be illustrated even better if the lower triangle is subtracted from the upper one (Fig. 10.2, right panel). The red (blue) colors show the squares in which more (less respectively) values are located. The more colorful the figure, the higher the asymmetry of the underlying copula. Not only the number of pairs in the upper left corner is increased, but also the number of pairs just below the bisecting line. These pairs

reflect the case where the rainfall values of the next time step are a little smaller compared to the current time step. The step wise reduction of the intensity illustrates the recession curve of an event.

Asymmetry of dependence In order to quantify the asymmetry one can evaluate the asymmetry function $A(k)$

$$A(k) = E[(U(t) - U(t + k))^3] \quad (10.4)$$

where $E[\cdot]$ denotes the expectation. Eq. 10.4 describes the differences in the copula density visualized in Fig. 10.2 (right panel). The differences to the power of three maintain the sign. The steep increase and long lasting declines, as in the case of precipitation, is indicated by a negative asymmetry. For the step wise calculation of the asymmetry as well as for the necessary modifications in the case of precipitation due to ties and non-unique values, I would like to refer to Müller *et al.* (2017a).

The asymmetry of a time series cannot be quantified using absolute values because the quantity depends on the autocorrelation. For low correlations, even small asymmetries can be significant. In order to evaluate the significance, the asymmetry values of an observed time series can be compared to asymmetry values of a precipitation time series based on an autoregressive process (AR). As the asymmetry of an AR is by definition zero, any calculated asymmetry in such time series is caused by chance. Consequently, a confidence interval of significant asymmetry can be defined using a large number of generated time series based on the AR. For the detailed explanation of the AR and the subsequent quantile-quantile-transformation to preserve the marginals, I would like to refer again to Müller *et al.* (2017a).

10.2. Data analysis using asymmetry of dependence

I analyzed the asymmetry of dependence of the location Augsburg and Kempten for different seasons (winter: September – April, summer: May – August) and different temporal aggregations (5 min – 24 h). Fig. 10.3 shows additionally the confidence intervals of precipitation time series based on an AR. Asymmetries outside the bandwidth indicate a significant asymmetry.

The time series of both stations exhibit significant asymmetric characteristics in the summer for all resolutions higher than daily aggregation. In winter the asymmetry is less pronounced and only significant for high resolutions. This indicates that the asymmetry is mainly related to short and heavy convective events. Furthermore, the asymmetry is a general characteristic and can not only be observed in Kempten and Augsburg, but also for several other stations in Bavaria and Singapore (Müller *et al.*, 2017a).

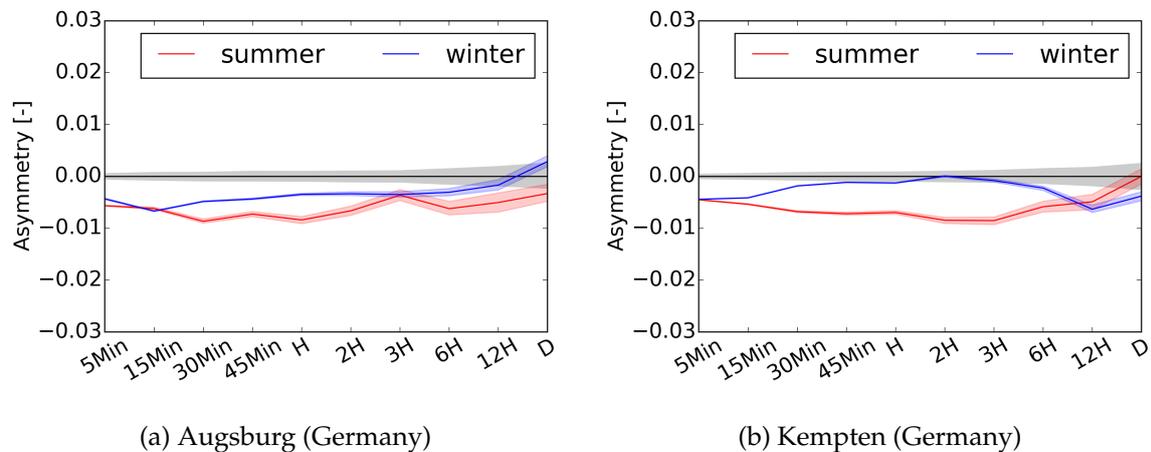


Figure 10.3.: Mean asymmetry $A(1)$ for different temporal aggregations and seasons. The grey confidence interval indicates the asymmetry of an AR(1) with the same correlation and marginal distribution as the measured time series. The colored bandwidth is caused by the ambiguous ranks of small values (figure from Müller *et al.*, 2017a).

10.3. Influence of asymmetry of dependence on combined sewers

The major question for my study is if asymmetric time series affect the results of sewer flows. As mentioned before, the effect of individual events on design tasks has already been investigated in literature. However, I could not find any analysis of influences of the irreversibility on long term CSO statistics. Therefore, I will investigate such influences using the artificial sewer introduced in Chap. 6.

In order to illustrate the effect of irreversibility, I reversed the original time series of Kempten and Augsburg previous to the sewer simulation. The CSO results of the temporal reversed time series is then compared with the sewer results using the original time series. In Fig. 10.4 the mean relative deviations of the CSO volume and duration is plotted. For overflow structures with storage (11 – 35) one can notice a slight overestimation of the volume and an underestimation of the duration. For structures without storage (R1 – R6), both statistics are overestimated.

The differences are in the same order as the remaining deviations of time series from the NiedSim generation (Fig. 8.15) and disaggregation (Fig. 9.7). Although the asymmetry is not considered in the statistics used for the NiedSim algorithm, it cannot explain most of the remaining deviations. Except for the CSO duration of the NiedSim generation, all other deviations are in the opposite direction. Consequently, the errors are not due to the missing asymmetry, but must be caused by other deficiencies in the time series.

For the analyzed sewer system the asymmetry seems to be less relevant and negligible compared to other uncertainties in precipitation modeling of NiedSim. Nevertheless, the

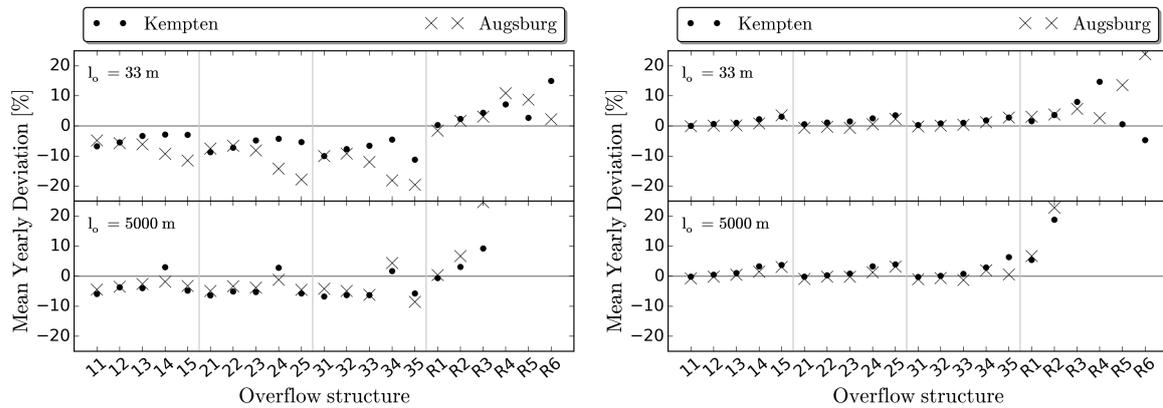


Figure 10.4.: CSO duration (left) and volume (right) for 21 different overflow structures and two different overland flow path lengths l_0 . The plots illustrate the relative deviation of the system response to the reverse a time series with respect to the original time series (figure from Müller *et al.*, 2017a).

asymmetry is a property that definitely influences combined sewers and can, for example, lead to reduced effects of sewer control (Müller *et al.*, 2017a). The more detailed sewer models and their simulations will become in future, the more realistic the rainfall models need to reflect the property of asymmetry.

11. Conclusions and Outlook

In my study, I showed that precipitation time series are highly complex and several statistics (global, event-based and extreme value statistics) must be regarded in order to validate the temporal structure of rainfall. Additionally, I illustrated the importance of an extensive analysis of observed precipitation data as an essential part of rainfall modeling. On the one hand, the investigation of measured data helps to benchmark synthetic time series by defining limits related to natural uncertainty. On the other hand, it can show artifacts caused by different measurement devices that can influence the optimization systematically. I also demonstrated large uncertainties of measured time series that are usually neglected in practice.

Synthetic precipitation time series cannot just be validated by comparing the target statistics of the optimization with the statistics of the simulated time series due to the complexity of rainfall. Instead, one must apply reasonable indirect validation methods that account for applications in urban hydrology. I judged the applicability of the synthetic time series for two main fields. First, for the hydraulic design that can be validated by partial series and second, for combined sewer simulations for which an artificial combined sewer model was used representing existing sewers in practice.

An inference from input precipitation statistics on the partial series or combined sewer model statistics is possible, but only for individual questions, for example, to detect errors in synthetic time series and to tweak the algorithm for their generation. A general conclusion about the importance of certain statistics for applications would be desirable, but cannot be given due to the complex rainfall characteristics and rainfall-runoff-overflow processes.

I illustrated certain deficiencies of synthetic time series of NiedSim2.x by an indirect validation. The problems could be related to difficulties with the optimization of the temporal structure at the hourly and 5 min aggregation. I showed that the hourly optimization is the crucial part with respect to most combined sewers. An issue of the hourly generation was the clustering of large precipitation values, whereas small and medium events were highly scattered and formed too many very short spells.

The optimization of the hourly time series is improved considerably by the positioning of large values with a Poisson process in advance of the optimization as well as by using wet spell frequencies as additional target in the objective function. Deviations in the most important statistics for hydraulic design and combined sewer simulations, that is in the partial series and in the mean annual CSO volume, could be systematically reduced. The remaining differences are in the range of the sampling uncertainty and, thus, seem to be tolerable for practical applications. For the mean annual CSO duration, however, larger deviations can still occur depending on the regarded sewer setup. In particular, for sewers

that show a high CSO activity, systematic underestimations may appear. The issues shown by the duration are related to deficiencies in the temporal structure of medium and small precipitation values. Unfortunately, I could neither refer them to global statistics nor to event-based statistics.

I demonstrated deficiencies in the analysis of the disaggregation caused by an erroneous temporal structure. The reason here is that 5 min values are spread across the entire hour instead of forming spells with a well-defined start and end. An optimization of the entropy at the 5 min aggregation could not improve the 5 min entropy for the right reasons and thus, had no positive effect on the temporal structure. The optimization using additional parameters seems to be necessary for a realistic representation of a high resolution temporal structure. However, in contrast to the hourly optimization, even one individual statistic is difficult to be optimized in the disaggregation. Consequently, a proper disaggregation using the approach of NiedSim seems to be challenging in general. Nevertheless, the results of the disaggregated time series are better compared to a series with a random intra-hourly structure.

An additional aspect makes the disaggregation difficult, that is, the high influence of the measurement device. I showed that different devices can have significant impacts on the observed high resolution statistics and thus, on the precipitation simulation. There is also some evidence that extremes could be affected as well. Consequently, high resolution long-term observations are inhomogeneous and must be cut into parts reflecting homogeneous statistics to be used for NiedSim. The observations of the most realistic measurement device cover only a relatively short time period. If only statistics were affected describing the temporal structure, the issue would be less problematic because these statistics are rather robust with respect to the annual variability. However, the reduced length of the time series is crucial for partial time series as they are highly vulnerable to sampling errors. Consequently, an indirect validation of the disaggregation with the partial series of measured data is problematic.

Extremes are not only crucial in the validation with the partial series of the disaggregation, but also for their optimization in the generation and disaggregation of NiedSim. Due to their vulnerability to sampling errors, a large bandwidth in observed extremes is possible. Consequently, the optimization of extremes using NiedSim introduces a large variability that might also affect the optimization of other statistics.

Extremes, however, cannot simply be ignored in the simulation of rainfall time series. Although they are not too important compared to other uncertainties for many combined sewer simulations, systematic deviations of extremes can lead to biases for sensitive sewer systems. Furthermore, extremes are very relevant for applications related to hydraulic design.

An approach to make the optimization of the temporal structure more robust is the non-consideration of extremes in the objective function. I showed in the study that the IDF-curves of synthetic time series are only slightly underestimated for most return periods. The remaining deviations could be adjusted in a post-processing step. As the partial series reflect only a very small fraction of all values of the precipitation time series, the general

temporal structure is not influenced much. Furthermore, if the values of the partial series are changed relatively to each other, the temporal structure between the adjusted values can be maintained. I analyzed a synthetic time series with an adjusted partial series. I did not show the results explicitly here, but I could neither find any considerable influence of such a post-processing approach on global or event-based statistics, nor on the CSO statistics of the combined sewer simulations.

Although extremes of synthetic time series are typically validated with IDF-curves in literature, I would like to mention a critical point here. The assumption of such a validation is that IDF-curves can fully describe the temporal structure of extremes. However, this is not the case as IDF-curves do not describe the internal temporal structure of extremes. Consequently, a perfect adjustment of the partial series in a post-processing step needs not necessarily reflect the extremes perfectly and, thus, might have some influence on further applications. Consequently, surcharge and flooding simulation with real case sewers should be performed in future to validate such a post-processing step.

Another advantage of such a method is that the extreme value analysis is independent of the generation of long-term precipitation time series. Consequently, different assumptions regarding the extreme value distributions could be applied to synthetic time series, for example, the widely used *Koordinierte Starkniederschlagsregionalisierung – Auswertung* [analysis of the coordinated regionalization of intense precipitation] (KOSTRA) extreme value data base, although the new update of KOSTRA2010 is discussed controversially (*Ihringer and Helms, 2016*).

My study is not only the basis for the enhanced algorithm of NiedSim3 (*Müller et al., 2017b*), but can also be the source for further investigations in the future. As already summarized, the reconstruction of well-defined spells in the disaggregation seems to be difficult for the NiedSim algorithm. Nevertheless, one could try to improve it by an optimization on the entropy for additional intra-hourly time scales such that not only the distribution, but also the temporal structure is reflected. Alternatively, one could try to optimize on wet spell distributions like in the hourly optimization.

As shown in my study, the temporal structure can be quite similar for certain locations. It could be worth to analyze the temporal structure of nearest neighbors in order to define representative structures for specific climates, regions or topographies. Such structures could then be used for a q-q-transformation as an alternative to the disaggregation of NiedSim. That is, a synthetic hourly time series could be q-q-transformed on the structure of representative high resolution observations.

A point I investigated only briefly was the inter-annual uncertainty of observed statistics and their influence on applications. Although the inter-annual variation of the temporal structure is not very distinct, some further investigations could be done here.

Some statistics like the extreme values cannot be optimized very well on an annual base. Due to the increasing computational power, the optimization of the time series needs not to be restricted to an annual base anymore. Consequently, one could include extremes in the objective function while optimizing on longer time periods.

With an increased demand of accuracy of applications, synthetic precipitation time series need to get more and more detailed. An important characteristic is the irreversibility of rainfall that is not considered in NiedSim so far. In future, one could implement methods based on the asymmetry of dependence to incorporate this property.

Another aspect is the investigation of the synthetic time series with respect to event-based CSO statistics instead of annual means. For example, for sewer management and RCT, realistic high resolution time series of single events are necessary to evaluate the performance of the flow control algorithm. Although measured real time data is usually used for this purpose, the synthetic data could be applied for long-term case studies. The correct representation of the annual means, as used in my study, does not necessarily reflect the correct event distribution. The observable deviations in the mean annual CSO duration shown in my study indicate potential issues in such event distributions. However, such investigations are challenging due to necessary event separation criteria.

I investigated the applicability of synthetic time series for sewer simulations in this study. However, several other fields of urban hydrology exist for which the performance of synthetic precipitation time series could be investigated, for example, low flow hydrology, catchment hydrology, LID approaches, urban water balance models and sewer system management.

The optimization of the temporal structure was developed based on locations of Southern Germany. As I have already demonstrated different performances here, where the climates are relatively similar, the conclusions of my work may not be transferred to other locations in the world. Therefore, the application of NiedSim to other climates would also be interesting for future research.

Bibliography

- Andrés-Doménech, I., A. Montanari, and J. B. Marco (2010), Stochastic rainfall analysis for storm tank performance evaluation, *Hydrology and Earth System Sciences*, 14(7), 1221–1232, doi:<http://dx.doi.org/10.5194/hess-14-1221-2010>.
- Aronica, G., G. Freni, and E. Oliveri (2005), Uncertainty analysis of the influence of rainfall time resolution in the modelling of urban drainage systems, *Hydrological Processes*, 19(5), 1055–1071, doi:<http://dx.doi.org/10.1002/hyp.5645>.
- Ashley, R., T. Hvitved-Jacobsen, and J.-L. Bertrand-Krajewski (1999), Quo vadis sewer process modelling?, *Water Science and Technology*, 39(9), 9 – 22, doi:[http://dx.doi.org/10.1016/S0273-1223\(99\)00211-5](http://dx.doi.org/10.1016/S0273-1223(99)00211-5).
- ATV-A128 (1992), *Richtlinien für die Bemessung und Gestaltung von Regenentlastungsanlagen in Mischwasserkanälen*, DWA, ISBN: 978-3-933693-16-7.
- ATV-A136 (1985), *Niederschlag, Aufbereitung und Weitergabe von Niederschlagsregistrierungen*, DWA, ISBN: 978-3-933693-21-1.
- ATV-A166 (2013), *Bauwerke der zentralen Regenwasserbehandlung und -rückhaltung - Konstruktive Gestaltung und Ausrüstung*, DWA, ISBN: 978-3-942964-50-0.
- ATV-DVWK-M165 (2004), *Anforderungen an Niederschlag-Abfluss-Berechnungen in der Siedlungsentwässerung*, DWA, ISBN: 978-3-924063-93-1.
- B. Efron, R. T. (1986), Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, *Statistical Science*, 1(1), 54–75.
- Bach, P. M., W. Rauch, P. S. Mikkelsen, D. T. McCarthy, and A. Deletic (2014), A critical review of integrated urban water modelling - urban drainage and beyond, *Environmental Modelling & Software*, 54, 88 – 107, doi:[10.1016/j.envsoft.2013.12.018](https://doi.org/10.1016/j.envsoft.2013.12.018).
- Barco, J., S. Papiri, and M. K. Stenstrom (2008), First flush in a combined sewer system, *Chemosphere*, 71(5), 827 – 833, doi:<http://dx.doi.org/10.1016/j.chemosphere.2007.11.049>.
- Bárdossy, A. (1993), Stochastische Modelle zur Beschreibung der raum-zeitlichen Variabilität des Niederschlags, Ph.D. thesis, Universität Karlsruhe (TH).
- Bárdossy, A. (1998), Generating precipitation time series using simulated annealing, *Water Resources Research*, 34(7), 1737–1744, doi:<http://dx.doi.org/10.1029/98WR00981>.

- Bárdossy, A., and G. Pegram (2014), Infilling missing precipitation records – a comparison of a new copula-based method with other techniques, *Journal of Hydrology*, 519, Part A, 1162 – 1170, doi:<https://doi.org/10.1016/j.jhydrol.2014.08.025>.
- Bárdossy, A., and E. J. Plate (1992), Space-time model for daily rainfall using atmospheric circulation patterns, *Water Resources Research*, 28(5), 1247–1259, doi:10.1029/91WR02589.
- Bárdossy, A., H. Giese, B. Haller, and J. Ruf (2000), Erzeugung synthetischer Niederschlagszeitreihen in hoher zeitlicher Auflösung für Baden-Württemberg, *Wasserwirtschaft*, 90, 548–553.
- Bárdossy, A., T. Müller, T. Mosthaf, and J. Seidel (2015), *Handbuch Niederschlags-Simulator 2.6*, Universität Stuttgart - IWS.
- Bartels, H., B. Dietzer, G. Malitz, F. Albrecht, and J. Guttenberger (2005), KOSTRA-DWD-2000–Starkniederschlagshöhen für Deutschland (1951-2000)–Fortschreibungsbericht, *Offenbach: Deutscher Wetterdienst, Hydrometeorologie*.
- Beck, F. (2012), Generation of spatially correlated synthetic rainfall time series in high temporal resolution - a data driven approach, Ph.D. thesis, University of Stuttgart, ISBN: 978-3-942036-23-8.
- Behrendt, J., E. Penda, A. Finkler, U. Heil, and C. Polte-Rudolf (2011), *Beschreibung der Datenbasis des NKDZ*, Deutscher Wetterdienst.
- Berger, C., and C. Falk (2009), Zustand der Kanalisation in Deutschland - Ergebnisse der DWA-Umfrage 2009, *DWA - Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V.*, pp. 1–15.
- Bernardara, P., C. D. Michele, and R. Rosso (2007), A simple model of rain in time: An alternating renewal process of wet and dry states with a fractional (non-gaussian) rain intensity, *Atmospheric Research*, 84(4), 291 – 301, doi:<http://dx.doi.org/10.1016/j.atmosres.2006.09.001>.
- Berne, A., G. Delrieu, J.-D. Creutin, and C. Obled (2004), Temporal and spatial resolution of rainfall measurements required for urban hydrology, *Journal of Hydrology*, 299(3–4), 166 – 179, doi:<http://dx.doi.org/10.1016/j.jhydrol.2004.08.002>, urban Hydrology.
- Beven, K. (2012), *Rainfall-Runoff Modelling: The Primer*, 2 ed., Wiley-Blackwell, Chichester, doi:<http://dx.doi.org/10.1002/9781119951001>.
- Brockerhoff, J. (2015), Untersuchung der Anwendbarkeit von synthetischen Niederschlagszeitreihen in der Schmutzfrachtsimulation mit SWMM, BSc. thesis - University of Stuttgart, supervisor: A. Bárdossy, advisor: T. Müller.
- Brommundt, J. (2008), Stochastische Generierung räumlich zusammenhängender Niederschlagszeitreihen, Ph.D. thesis, Universität Stuttgart, ISBN: 3-933761-74-3.
- Brunner, P. G. (1997), *Wirtschaftliche Aspekte bei Gestaltung, Konstruktion und Ausrüstung von Regenbecken*, 148 S. pp., LfU, Referat 15 - Bibliothek, Karlsruhe.

- Burgueño, A., E. Vilar, and M. Puigcerver (1990), Spectral analysis of 49 years of rainfall rate and relation to fade dynamics, *IEEE Transactions on Communications*, 38(9), 1359–1366, doi:10.1109/26.61377.
- Burian, S. J., and F. G. Edwards (2002), Historical perspectives of urban drainage, in *9th International Conference on Urban Drainage*, pp. 1–16, doi:http://dx.doi.org/10.1061/40644(2002)284.
- Burian, S. J., S. R. Durrans, S. Tomic, R. L. Pimmel, and C. N. Wai (2000), Rainfall disaggregation using artificial neural networks, *Journal of Hydrologic Engineering*, 5(3), 299–307, doi:http://dx.doi.org/10.1061/(ASCE)1084-0699(2000)5:3(299).
- Burlando, P., and R. Rosso (1996), Scaling and multiscaling models of depth-duration-frequency curves for storm precipitation, *Journal of Hydrology*, 187(1), 45 – 64, doi:http://dx.doi.org/10.1016/S0022-1694(96)03086-7.
- Butler, D., and J. Davies (2011), *Urban drainage*, 3 ed., Spon Press, ISBN: 0-419-22340-1.
- BWK-M3 (2007), *Ableitung von immissionsorientierten Anforderungen an Misch- und Niederschlagswassereinleitungen unter Berücksichtigung örtlicher Verhältnisse*, Bund der Ingenieure für Wasserwirtschaft, Abfallwirtschaft und Kulturbau (BWK) e.V., ISBN: 978-3-8167-7471-6.
- Chandler, R. E., and H. S. Wheeler (2002), Analysis of rainfall variability using generalized linear models: A case study from the west of ireland, *Water Resources Research*, 38(10), 10–1–10–11, doi:http://dx.doi.org/10.1029/2001WR000906, 1192.
- Cowpertwait, P. S. P. (1995), A generalized spatial-temporal model of rainfall based on a clustered point process, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 450(1938), 163–175, doi:http://dx.doi.org/10.1098/rspa.1995.0077.
- Cunnane, C. (1978), Unbiased plotting positions - a review, *Journal of Hydrology*, 37(3), 205 – 222, doi:10.1016/0022-1694(78)90017-3.
- Dolšak, D., N. Bezak, and M. Šraj (2016), Temporal characteristics of rainfall events under three climate types in slovenia, *Journal of Hydrology*, 541, Part B, 1395 – 1405, doi:http://dx.doi.org/10.1016/j.jhydrol.2016.08.047.
- Dooge, J. C. (1959), A general theory of the unit hydrograph, *Journal of Geophysical Research*, 64(2), 241–256, doi:http://dx.doi.org/10.1029/JZ064i002p00241.
- Drechsel, U. (1991), Repräsentanz und Übertragbarkeit von Niederschlagsersatzbelastungen zur Durchführung von Schmutzfrachtberechnungen, *Technische Berichte über Ingenieurhydrology und Hydraulik*, 45, 1–112.
- Dunkerley, D. (2008a), Rain event properties in nature and in rainfall simulation experiments: a comparative review with recommendations for increasingly systematic study and reporting, *Hydrological Processes*, 22(22), 4415–4435, doi:http://dx.doi.org/10.1002/hyp.7045.

- Dunkerley, D. (2008b), Identifying individual rain events from pluviograph records: a review with analysis of data from an Australian dryland site, *Hydrological Processes*, 22(26), 5024–5036, doi:<http://dx.doi.org/10.1002/hyp.7122>.
- DWA-A100 (2006), *Leitlinien der integralen Siedlungsentwässerung (ISiE)*, DWA, ISBN: 978-3-939057-70-3.
- DWA-A102 (2016), *Grundsätze zur Bewirtschaftung und Behandlung von Regenwasserabflüssen zur Einleitung in Oberflächengewässer*, DWA, Entwurf.
- DWA-A118 (2006), *Hydraulische Bemessung und Nachweis von Entwässerungssystemen*, DWA, ISBN:978-3-939057-15-4.
- DWA-A531 (2012), *Starkregen in Abhängigkeit von Wiederkehrzeit und Dauer*, DWA, ISBN: 978-3-942964-28-9.
- DWA-M153 (2007), *Handlungsempfehlung zum Umgang mit Regenwasser*, DWA, ISBN: 978-3-939057-98-7.
- Elliott, A., and S. Trowsdale (2007), A review of models for low impact urban stormwater drainage, *Environmental Modelling & Software*, 22(3), 394–405, doi:10.1016/j.envsoft.2005.12.005, special section: Advanced Technology for Environmental Modelling.
- EU (2000), *Water Framework Directive. Directive 2000/60/EC European Parliament and Council*.
- Fletcher, T., H. Andrieu, and P. Hamel (2013), Understanding, management and modelling of urban hydrology and its consequences for receiving waters: A state of the art, *Advances in Water Resources*, 51, 261 – 279, doi:<http://dx.doi.org/10.1016/j.advwatres.2012.09.001>, 35th Year Anniversary Issue.
- Fraedrich, K., and C. Larnder (1993), Scaling regimes of composite rainfall time series, *Tellus A*, 45(4), 289–298, doi:<http://dx.doi.org/10.1034/j.1600-0870.1993.t01-3-00004.x>.
- García, L., J. Barreiro-Gomez, E. Escobar, D. Téllez, N. Quijano, and C. Ocampo-Martinez (2015), Modeling and real-time control of urban drainage systems: A review, *Advances in Water Resources*, 85, 120–132, doi:10.1016/j.advwatres.2015.08.007.
- Gaume, E., N. Mouhous, and H. Andrieu (2007), Rainfall stochastic disaggregation models: Calibration and validation of a multiplicative cascade model, *Advances in Water Resources*, 30(5), 1301 – 1319, doi:<http://dx.doi.org/10.1016/j.advwatres.2006.11.007>.
- Giannakis, G. B., and M. K. Tsatsanis (1994), Time-domain tests for gaussianity and time-reversibility, *IEEE Transactions on Signal Processing*, 42(12), 3460–3472, doi:<http://dx.doi.org/10.1109/78.340780>.
- Giesecke, J., and U. Haberlandt (1998), Precipitation data requirements for urban hydrology, *Water international*, 23(2), 60–66, doi:<http://dx.doi.org/10.1080/02508069808686743>.

- Göbel, P., C. Dierkes, and W. Coldewey (2007), Storm water runoff concentration matrix for urban areas, *Journal of Contaminant Hydrology*, 91(1–2), 26 – 42, doi:<http://dx.doi.org/10.1016/j.jconhyd.2006.08.008>, issues in urban hydrology: The emerging field of urban contaminant hydrology.
- Goovaerts, P. (2000), Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *Journal of Hydrology*, 228(1–2), 113 – 129, doi:[http://dx.doi.org/10.1016/S0022-1694\(00\)00144-X](http://dx.doi.org/10.1016/S0022-1694(00)00144-X).
- Gunzenhauser, S. (2016), Sensitivitätsanalyse der Disaggregation des stochastischen Niederschlagsgenerators NiedSim, BSc. thesis - University of Stuttgart, supervisor: A. Bárdossy, advisor: T. Müller.
- Haberlandt, U. (1996), Stochastische Synthese und Regionalisierung des Niederschlages für Schmutzfrachtberechnungen, Ph.D. thesis, Universität Stuttgart, ISBN: 3-921694-88-4.
- Haberlandt, U. (1998), Stochastic rainfall synthesis using regionalized model parameters, *Journal of hydrologic engineering*, 3(3), 160–168, doi:[http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(1998\)3%3A3\(160\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(1998)3%3A3(160)).
- Hammond, M., A. Chen, S. Djordjević, D. Butler, and O. Mark (2015), Urban flood impact assessment: A state-of-the-art review, *Urban Water Journal*, 12(1), 14–29, doi:[10.1080/1573062X.2013.857421](https://doi.org/10.1080/1573062X.2013.857421).
- Hingray, B., and M. B. Haha (2005), Statistical performances of various deterministic and stochastic models for rainfall series disaggregation, *Atmospheric Research*, 77(1–4), 152 – 175, doi:<http://dx.doi.org/10.1016/j.atmosres.2004.10.023>, precipitation in Urban Areas - 6th International Workshop on Precipitation in Urban Areas.
- Hoang, C. T., I. Tchiguirinskaia, D. Schertzer, P. Arnaud, J. Lavabre, and S. Lovejoy (2012), Assessing the high frequency quality of long rainfall series, *Journal of Hydrology*, 438, 39–51, doi:<http://dx.doi.org/10.1016/j.jhydrol.2012.01.044>.
- Huff, F. A. (1967), Time distribution of rainfall in heavy storms, *Water Resources Research*, 3(4), 1007–1019, doi:<http://dx.doi.org/10.1029/WR003i004p01007>.
- Ihringer, J., and M. Helms (2016), KOSTRA-DWD-2010 Bewertung im Hinblick auf die wasserwirtschaftliche Bemessungspraxis, LUBW - Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg.
- itwh (2009), *Kontinuierliche Simulation zur Bemessung von Speicherbauwerken in urbanen Entwässerungssystemen. KOSIM 7.3. Modellbeschreibung.*, Institut für technisch-wissenschaftliche Hydrologie GmbH, Hannover.
- Jacobson, C. R. (2011), Identification and quantification of the hydrological impacts of imperviousness in urban catchments: A review, *Journal of Environmental Management*, 92(6), 1438–1448, doi:<http://dx.doi.org/10.1016/j.jenvman.2011.01.018>.

- Jarraud, M. (2008), Guide to meteorological instruments and methods of observation (WMO-No. 8), *World Meteorological Organisation: Geneva, Switzerland*, ISBN: 978-92-63-10008-5.
- Katz, R. W., and M. B. Parlange (1995), Generalizations of chain-dependent processes: Application to hourly precipitation, *Water Resources Research*, 31(EFLUM-ARTICLE-1995-007), 1331–1341, doi:http://dx.doi.org/10.1029/94WR03152.
- Koutsoyiannis, D., and C. Onof (2001), Rainfall disaggregation using adjusting procedures on a Poisson cluster model, *Journal of Hydrology*, 246(1–4), 109 – 122, doi:http://dx.doi.org/10.1016/S0022-1694(01)00363-8.
- Koutsoyiannis, D., D. Kozonis, and A. Manetas (1998), A mathematical framework for studying rainfall intensity-duration-frequency relationships, *Journal of Hydrology*, 206(1), 118–135, doi:http://dx.doi.org/10.1016/S0022-1694(98)00097-3.
- Krause, P., D. P. Boyle, and F. Bäse (2005), Comparison of different efficiency criteria for hydrological model assessment, *Advances in Geosciences*, 5, 89–97, doi:http://dx.doi.org/10.5194/adgeo-5-89-2005.
- Lanza, L., M. Leroy, C. Alexandropoulos, L. Stagi, and W. Wauben (2005), WMO Laboratory intercomparison of rainfall intensity gauges, *Instruments and Observing Methods Rep. WMO/TD-No. 1304*, World Meteorological Organization.
- Lau, J., D. Butler, and M. Schütze (2002), Is combined sewer overflow spill frequency/volume a good indicator of receiving water quality impact?, *Urban water*, 4(2), 181–189, doi:http://dx.doi.org/10.1016/S1462-0758(02)00013-4.
- Lawrance, A. J. (1991), Directionality and reversibility in time series, *International Statistical Review / Revue Internationale de Statistique*, 59(1), 67–79, doi:http://dx.doi.org/10.2307/1403575.
- Lombardo, F., E. Volpi, D. Koutsoyiannis, and S. Papalexiou (2014), Just two moments! A cautionary note against use of high-order moments in multifractal models in hydrology, *Hydrology and Earth System Sciences*, 18(1), 243–255, doi:http://dx.doi.org/10.5194/hess-18-243-2014.
- Makkonen, L. (2006), Plotting positions in extreme value analysis, *Journal of Applied Meteorology and Climatology*, 45(2), 334–340, doi:http://dx.doi.org/10.1175/JAM2349.1.
- Malitz, G., and H. Ertel (2015), KOSTRA-DWD-2010 Starkniederschlagshöhen für Deutschland (Bezugszeitraum 1951 bis 2010)-Abschlussbericht, *Offenbach: Deutscher Wetterdienst, Hydrometeorologie*.
- Marshall, A. W., and I. Olkin (2007), *Life distributions*, Springer.
- Müller, T., U. Dittmer, and A. Bárdossy (2014), Generierung von synthetischen Niederschlagszeitreihen für die Schmutzfrachtsimulation, in *4. Aqua Urbanica Konferenz. Misch- und Niederschlagswasserbehandlung im urbanen Raum*, 23. / 24. Oktober 2014, Innsbruck.

- Molini, A., G. G. Katul, and A. Porporato (2009), Revisiting rainfall clustering and intermittency across different climatic regimes, *Water resources research*, 45(11), doi:<http://dx.doi.org/10.1029/2008WR007352>.
- Molnar, P., and P. Burlando (2005), Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, *Atmospheric Research*, 77(1), 137–151, doi:<http://dx.doi.org/10.1016/j.atmosres.2004.10.024>.
- Mosthaf, T. (2017), New concepts for regionalizing temporal distributions of precipitation and for its application in spatial rainfall simulation, Ph.D. thesis, University of Stuttgart.
- Müller, H., and U. Haberlandt (2016), Temporal rainfall disaggregation using a multiplicative cascade model for spatial application in urban hydrology, *Journal of Hydrology*, pp. –, doi:<http://dx.doi.org/10.1016/j.jhydrol.2016.01.031>.
- Müller, T., M. Schütze, and A. Bárdossy (2017a), Temporal asymmetry in precipitation time series and its influence on flow simulations in combined sewer systems, *Advances in Water Resources*, doi:<http://dx.doi.org/10.1016/j.advwatres.2017.06.010>.
- Müller, T., T. Mosthaf, S. Gunzenhauser, J. Seidel, and A. Bárdossy (2017b), *Grundlagenbericht Niederschlags-Simulator (NiedSim3)*, Mitteilungen Institut für Wasser- und Umweltsystemmodellierung, Universität, H. 255, ISBN: 978-3-942036-59-7.
- NRSC (2004), *National Engineering Handbook, Part 630, Chapter 10: Estimation of Direct Runoff from Storm Rainfall*, Natural Resources Conservation Service.
- Nystuen, J. A., J. R. Proni, P. G. Black, and J. C. Wilkerson (1996), A comparison of automatic rain gauges, *Journal of Atmospheric and Oceanic Technology*, 13(1), 62–73, doi:[http://dx.doi.org/10.1175/1520-0426\(1996\)013<0062:ACOARG>2.0.CO;2](http://dx.doi.org/10.1175/1520-0426(1996)013<0062:ACOARG>2.0.CO;2).
- Ochoa-Rodriguez, S., L.-P. Wang, A. Gires, R. D. Pina, R. Reinoso-Rondinel, G. Bruni, A. Ichiba, S. Gaitan, E. Cristiano, J. van Assel, S. Kroll, D. Murlà-Tuyls, B. Tisserand, D. Schertzer, I. Tchiguirinskaia, C. Onof, P. Willems, and M.-C. ten Veldhuis (2015), Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation, *Journal of Hydrology*, 531, Part 2, 389 – 407, doi:<http://dx.doi.org/10.1016/j.jhydrol.2015.05.035>, hydrologic Applications of Weather Radar.
- Onof, C., E. R. Chandler, A. Kakou, P. Northrop, S. H. Wheeler, and V. Isham (2000), Rainfall modelling using poisson-cluster processes: a review of developments, *Stochastic Environmental Research and Risk Assessment*, 14(6), 384–411, doi:[10.1007/s004770000043](https://doi.org/10.1007/s004770000043).
- Onof, C., J. Townend, and R. Kee (2005), Comparison of two hourly to 5-min rainfall disaggregators, *Atmospheric Research*, 77(1–4), 176 – 187, doi:<http://dx.doi.org/10.1016/j.atmosres.2004.10.022>, Precipitation in Urban Areas, 6th International Workshop on Precipitation in Urban Areas.
- Ormsbee, L. E. (1989), Rainfall disaggregation model for continuous hydrologic modeling, *Journal of Hydraulic Engineering*, 115(4), 507–525, doi:[http://dx.doi.org/10.1061/\(ASCE\)0733-9429\(1989\)115:4\(507\)](http://dx.doi.org/10.1061/(ASCE)0733-9429(1989)115:4(507)).

- Paschalis, A., P. Molnar, S. Fatichi, and P. Burlando (2014), On temporal stochastic modeling of precipitation, nesting models across scales, *Advances in Water Resources*, 63, 152 – 166, doi:http://dx.doi.org/10.1016/j.advwatres.2013.11.006.
- Pui, A., A. Sharma, R. Mehrotra, B. Sivakumar, and E. Jeremiah (2012), A comparison of alternatives for daily to sub-daily rainfall disaggregation, *Journal of Hydrology*, 470–471, 138 – 157, doi:http://dx.doi.org/10.1016/j.jhydrol.2012.08.041.
- Python (2015), Python 2.7.10 documentation, <http://docs.python.org/2.7/>, accessed: 2015-06-29.
- Restrepo-Posada, P. J., and P. S. Eagleson (1982), Identification of independent rainstorms, *Journal of Hydrology*, 55(1), 303–319, doi:http://dx.doi.org/10.1016/0022-1694(82)90136-6.
- Robinson, J. S., and M. Sivapalan (1997), Temporal scales and hydrological regimes: Implications for flood frequency scaling, *Water Resources Research*, 33(12), 2981–2999, doi: <http://dx.doi.org/10.1029/97WR01964>.
- Rodriguez-Iturbe, I., D. Cox, and V. Isham (1987), Some models for rainfall based on stochastic point processes, *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 410(1839), 269–288, doi:http://dx.doi.org/10.1098/rspa.1987.0039.
- Rossman, L. A., and W. C. Huber (2016), *Storm Water Management Model Reference Manual Volume I - Hydrology (Revised)*, United States Environmental Protection Agency (EPA).
- Salvadori, G., and C. D. Michele (2006), Statistical characterization of temporal structure of storms, *Advances in Water Resources*, 29(6), 827 – 842, doi:http://dx.doi.org/10.1016/j.advwatres.2005.07.013.
- Schilling, W. (1991), Rainfall data for urban hydrology: what do we need?, *Atmospheric Research*, 27(1–3), 5 – 21, doi:http://dx.doi.org/10.1016/0169-8095(91)90003-F.
- Schleiss, M., and J. A. Smith (2016), Two simple metrics for quantifying rainfall intermittency: The burstiness and memory of interamount times, *Journal of Hydrometeorology*, 17(1), 421–436, doi:http://dx.doi.org/10.1175/JHM-D-15-0078.1.
- Schmit, T. G., and H. H. Hahn (1986), *Schmutzfrachtberechnung für Kanalisationsnetze*, ISWW Karlsruhe, vol. 44, Inst. für Siedlungswasserwirtschaft (ISWW), Univ. Karlsruhe.
- Schmitt, T. G., M. Thomas, and N. Ettrich (2004), Analysis and modeling of flooding in urban drainage systems, *Journal of Hydrology*, 299(3–4), 300 – 311, doi:http://dx.doi.org/10.1016/j.jhydrol.2004.08.012, urban Hydrology.
- Schütze, M., D. Butler, and B. M. Beck (2011), *Modelling, simulation and control of urban wastewater systems*, Springer Science & Business Media, doi:http://dx.doi.org/10.1007/978-1-4471-0157-4.
- Semadeni-Davies, A., C. Hernebring, G. Svensson, and L.-G. Gustafsson (2008), The impacts of climate change and urbanisation on drainage in Helsingborg, Sweden: Combined sewer system, *Journal of Hydrology*, 350(1–2), 100 – 113, doi:http://dx.doi.org/10.1016/j.jhydrol.2007.05.028.

- Sevruk, B., and W. Hamon (1984), International comparison of national precipitation gauges with a reference pit gauge, *Instruments and Observing Methods Rep. 17*, World Meteorological Organization.
- Sevruk, B., and S. Klemm (1989), Catalogue of national standard precipitation gauges, *Instruments and Observing Methods Rep. WMO/TD-No. 313*, World Meteorological Organization.
- Singh, V. P. (2011), Hydrologic synthesis using entropy theory: Review, *Journal of Hydrologic Engineering*, 16(5), 421–433, doi:http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000332.
- Steuber, T. L., P. C. Kiessler, and R. Lund (2012), Testing for reversibility in Markov chain data, *Probability in the Engineering and Informational Sciences*, 26(4), 593–611, doi:http://dx.doi.org/10.1017/S0269964812000228.
- Strangeways, I. (2006), *Precipitation: theory, measurement and distribution*, Cambridge University Press.
- Sugimoto, T., A. Bárdossy, G. G. S. Pegram, and J. Cullmann (2016), Investigation of hydrological time series using copulas for detecting catchment characteristics and anthropogenic impacts, *Hydrology and Earth System Sciences*, 20(7), 2705–2720, doi:http://dx.doi.org/10.5194/hess-20-2705-2016.
- van der Heijden, S., A. Callau, H. Müller, C. Berndt, T. Müller, T. Mosthaf, A. Wagner, M. Lorenz, S. Rohde, K. Schroeder, K. Teuber, S. Maßmann, M. Schönfeld, and S. Krämer (2017), Abschlussbericht für das INIS-Projekt SYNOPSE, Institut für Hydrologie und Wasserwirtschaft, Hannover.
- Vandenberghe, S., N. E. C. Verhoest, C. Onof, and B. De Baets (2011), A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall, *Water Resources Research*, 47(7), n/a–n/a, doi:http://dx.doi.org/10.1029/2009WR008388, w07529.
- Verhoest, N., P. A. Troch, and F. P. D. Troch (1997), On the applicability of bartlett-lewis rectangular pulses models in the modeling of design storms at a point, *Journal of Hydrology*, 202(1–4), 108 – 120, doi:10.1016/S0022-1694(97)00060-7.
- Wackernagel, H. (1995), *Multivariate geostatistics: An introduction with applications*, 1 ed., Springer-Verlag Berlin Heidelberg, doi:http://dx.doi.org/10.1061/40644(2002)284, ISBN: 3-540-60127-9.
- Westra, S., R. Mehrotra, A. Sharma, and R. Srikanthan (2012), Continuous rainfall simulation: 1. A regionalized subdaily disaggregation approach, *Water Resources Research*, 48(1), doi:http://dx.doi.org/10.1029/2011WR010489.
- Wilks, D. S. (2011), *Statistical methods in the atmospheric sciences*, vol. 100, Academic press, ISBN: 978-0-12-385022-5.

- Wilks, D. S., and R. L. Wilby (1999), The weather generation game: a review of stochastic weather models, *Progress in Physical Geography*, 23(3), 329–357, doi:http://dx.doi.org/10.1177/030913339902300302.
- Xanthopoulos, C. (1990), *Methode für die Entwicklung von Modellregenspektren für die Schmutzfrachtberechnung*, ISWW Karlsruhe, vol. 57, Inst. für Siedlungswasserwirtschaft (ISWW), Univ. Karlsruhe.
- Young, C. B., B. M. McEnroe, and A. C. Rome (2009), Empirical determination of rational method runoff coefficients, *Journal of Hydrologic Engineering*, 14(12), 1283–1289, doi:http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000114.
- Yu, R., W. Yuan, and J. Li (2013), The asymmetry of rainfall process, *Chinese Science Bulletin*, 58(16), 1850–1856, doi:http://dx.doi.org/10.1007/s11434-012-5653-6.
- Zhu, H.-j., and W. Schilling (1996), Simulation errors due to insufficient temporal rainfall resolution - annual combined sewer overflow, *Atmospheric research*, 42(1), 19–32, doi:10.1016/0169-8095(95)00050-X.
- Zoppou, C. (2001), Review of urban storm water models, *Environmental Modelling & Software*, 16(3), 195 – 231, doi:http://dx.doi.org/10.1016/S1364-8152(00)00084-0.

A. Appendix

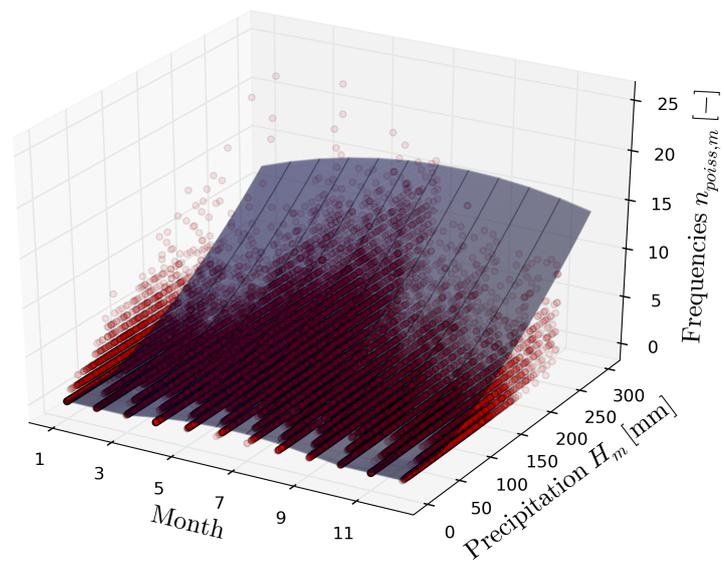


Figure A.1.: Least-square optimization of the 3d-surface between months, monthly precipitation sums and number values of the Poisson process ($x_{poiss} > 4$ mm).

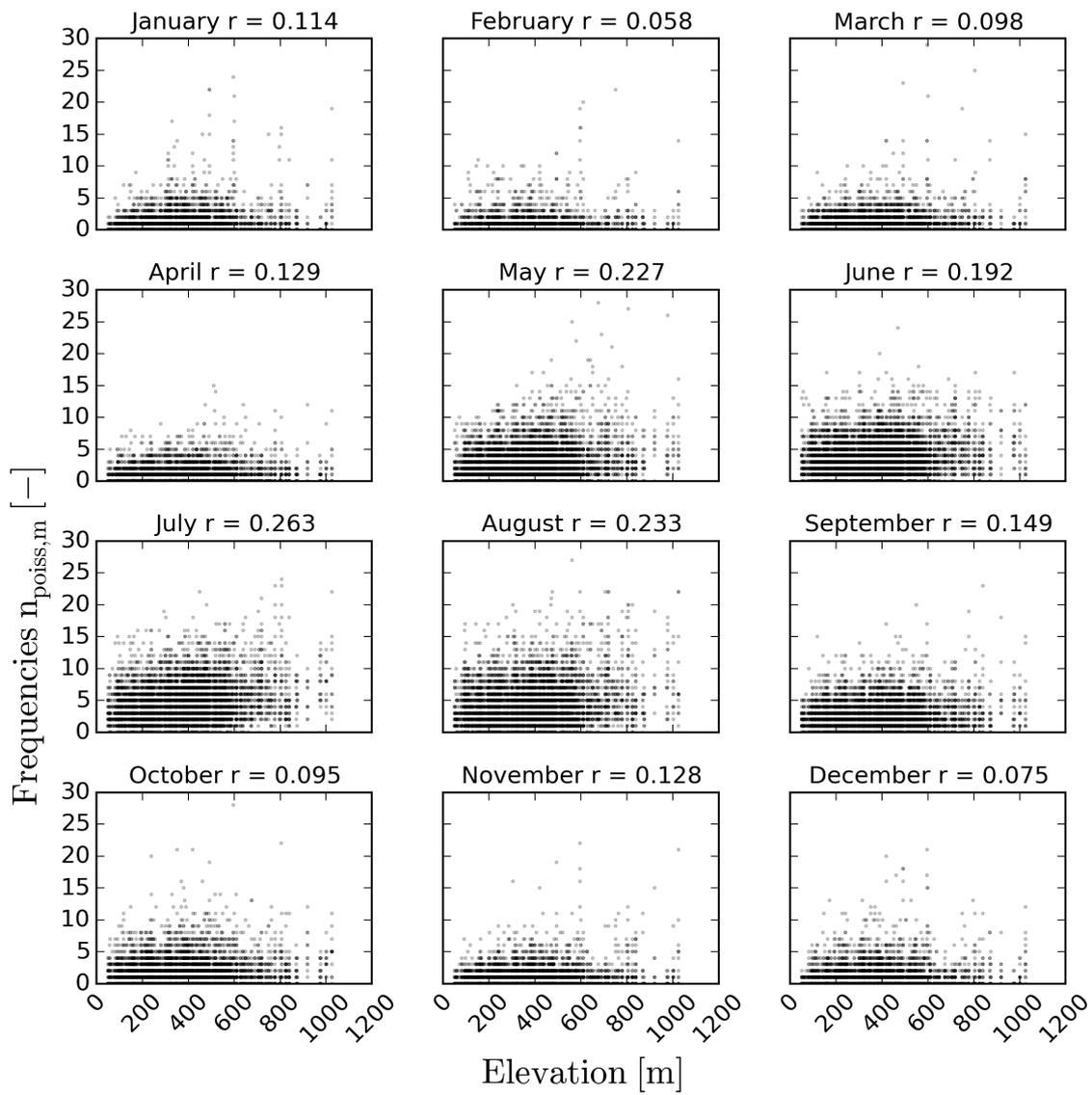


Figure A.2.: Scatterplot between elevation and monthly frequency of the Poisson process ($x_{\text{poiss}} > 4$ mm). r denotes the correlation coefficient of the linear regression.

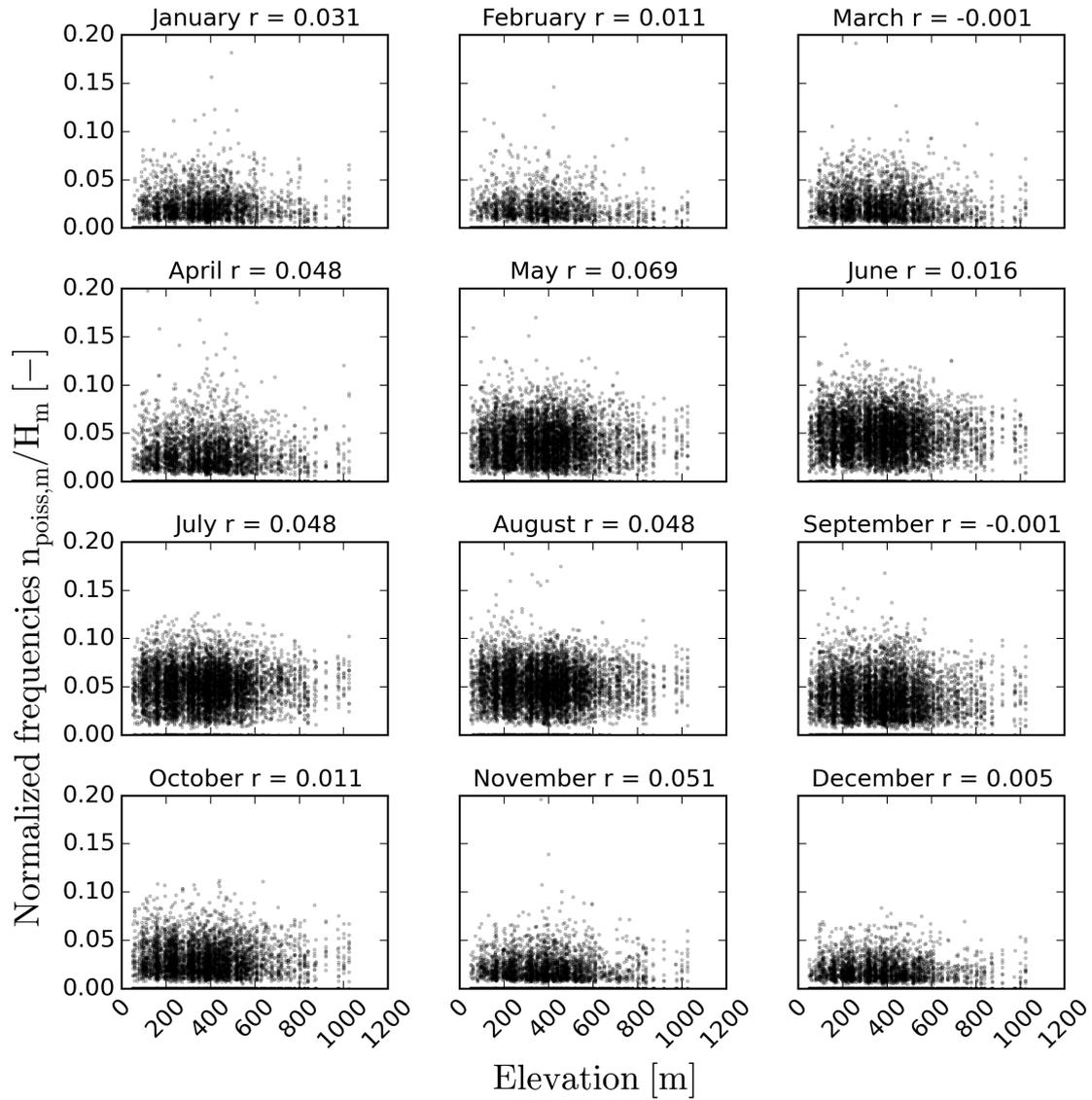


Figure A.3.: Scatterplot between elevation and monthly frequency of the Poisson process normalized with the monthly precipitation sum ($x_{\text{poiss}} > 4$ mm). r denotes the correlation coefficient of the linear regression

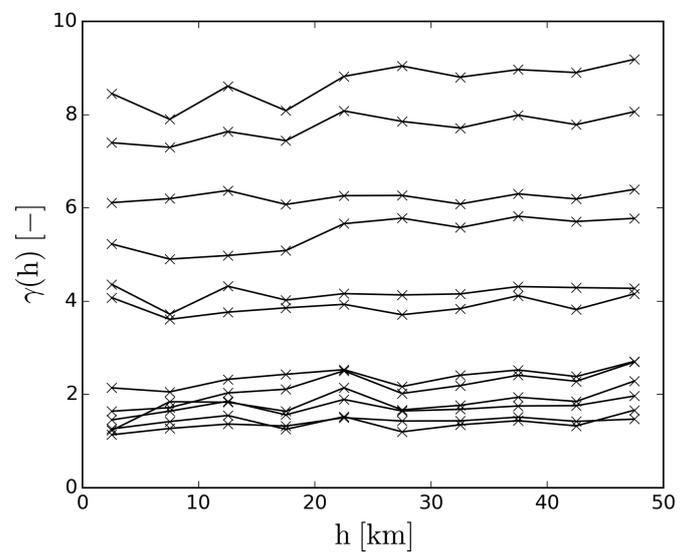
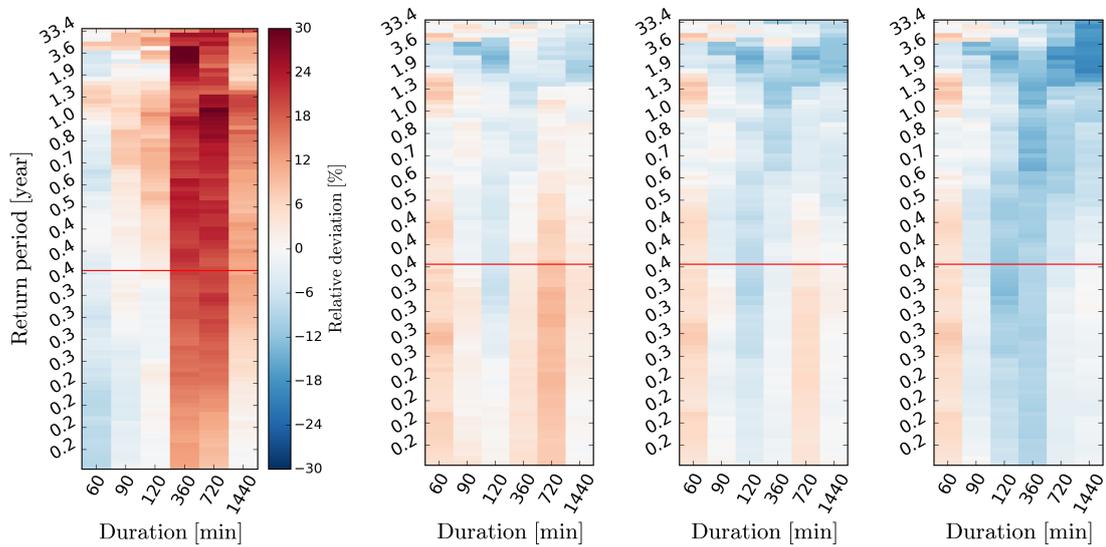


Figure A.4.: Semivariogram of the Poisson values of different months with distance classes of 5 km.

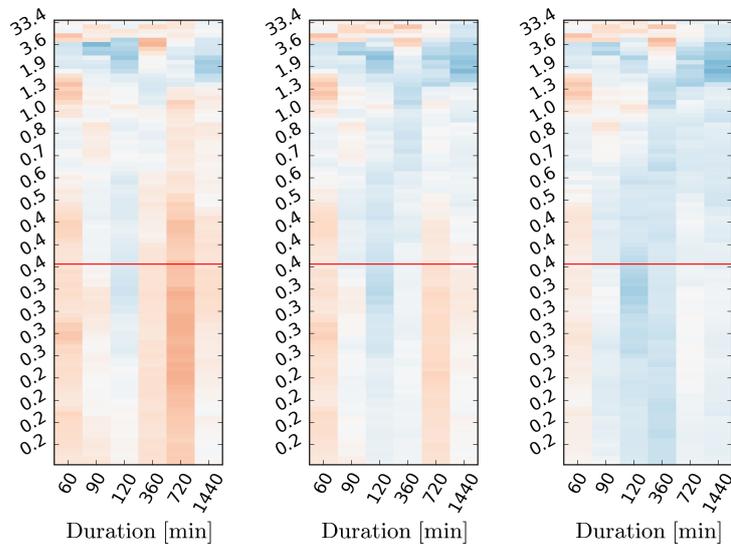


(a) Version: a

(b) Version: p5

(c) Version: p4

(d) Version: p3



(e) Version: p5x

(f) Version: p4x

(g) Version: p3x

Figure A.5.: Relative deviation of partial series of different versions of hourly optimized time series from the original time series. Return periods above the red line are usually considered for extreme value analysis. All data refer to Kempten, 1961 - 1980.

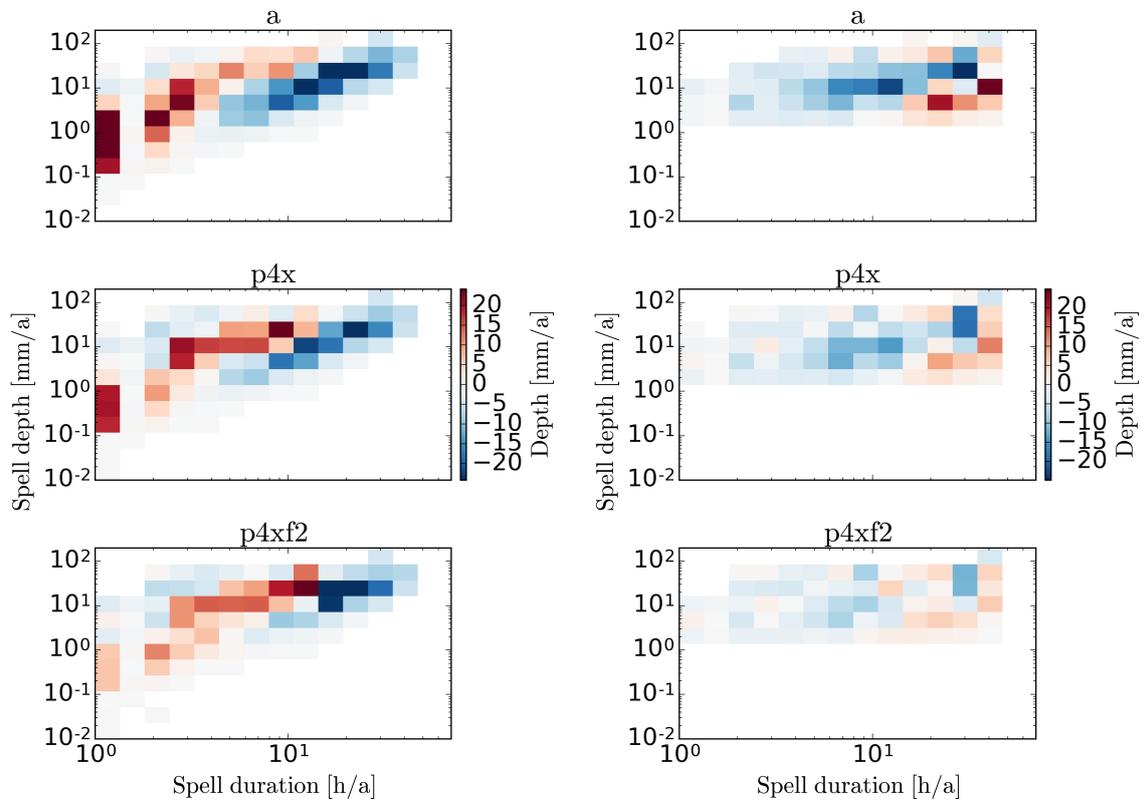
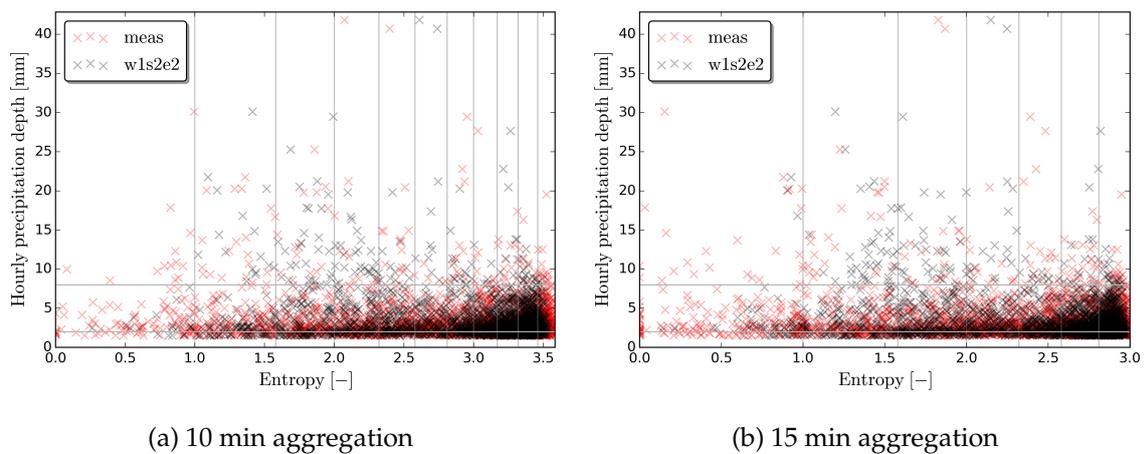


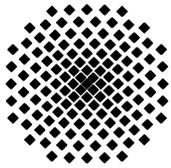
Figure A.6.: Dependence structure of precipitation spells separated with $t_{se} = 1$ h, $i_e = 0$ mm, $d_e = 0$ mm (left) and $t_{se} = 8$ h, $i_e = 0,1$ mm, $d_e = 1,5$ mm (right). Both figures show the cumulative precipitation depth. For a description see text or Fig. 8.14.



(a) 10 min aggregation

(b) 15 min aggregation

Figure A.7.: Scatter plot of entropies (calculated at 10 and 15 min aggregation) and hourly precipitation sum.



Institut für Wasser- und Umweltsystemmodellierung Universität Stuttgart

Pfaffenwaldring 61
70569 Stuttgart (Vaihingen)
Telefon (0711) 685 - 64717/64749/64752/64679
Telefax (0711) 685 - 67020 o. 64746 o. 64681
E-Mail: iws@iws.uni-stuttgart.de
<http://www.iws.uni-stuttgart.de>

Direktoren

Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy
Prof. Dr.-Ing. Rainer Helmig
Prof. Dr.-Ing. Silke Wieprecht
Prof. Dr.-Ing. Wolfgang Nowak

Vorstand (Stand 1.3.2017)

Prof. Dr. rer. nat. Dr.-Ing. A. Bárdossy
Prof. Dr.-Ing. R. Helmig
Prof. Dr.-Ing. S. Wieprecht
Prof. Dr. J.A. Sander Huisman
Jürgen Braun, PhD
apl. Prof. Dr.-Ing. H. Class
Dr.-Ing. H.-P. Koschitzky
Dr.-Ing. M. Noack
Prof. Dr.-Ing. W. Nowak
Dr. rer. nat. J. Seidel
Dr.-Ing. K. Terheiden
Dr.-Ing. habil. Sergey Oladyshkin

Emeriti

Prof. Dr.-Ing. habil. Dr.-Ing. E.h. Jürgen Giesecke
Prof. Dr.h.c. Dr.-Ing. E.h. Helmut Kobus, PhD

Lehrstuhl für Wasserbau und Wassermengenwirtschaft

Leiter: Prof. Dr.-Ing. Silke Wieprecht
Stellv.: Dr.-Ing. Kristina Terheiden
Versuchsanstalt für Wasserbau
Leiter: Dr.-Ing. Markus Noack

Lehrstuhl für Hydromechanik und Hydrosystemmodellierung

Leiter: Prof. Dr.-Ing. Rainer Helmig
Stellv.: apl. Prof. Dr.-Ing. Holger Class

Lehrstuhl für Hydrologie und Geohydrologie

Leiter: Prof. Dr. rer. nat. Dr.-Ing. András Bárdossy
Stellv.: Dr. rer. nat. Jochen Seidel
Hydrogeophysik der Vadosen Zone
(mit Forschungszentrum Jülich)
Leiter: Prof. Dr. J.A. Sander Huisman

Lehrstuhl für Stochastische Simulation und Sicherheitsforschung für Hydrosysteme

Leiter: Prof. Dr.-Ing. Wolfgang Nowak
Stellv.: Dr.-Ing. habil. Sergey Oladyshkin

VEGAS, Versuchseinrichtung zur Grundwasser- und Altlastensanierung

Leitung: Jürgen Braun, PhD, AD
Dr.-Ing. Hans-Peter Koschitzky, AD

Verzeichnis der Mitteilungshefte

- 1 Röhnisch, Arthur: *Die Bemühungen um eine Wasserbauliche Versuchsanstalt an der Technischen Hochschule Stuttgart*, und Fattah Abouleid, Abdel: *Beitrag zur Berechnung einer in lockeren Sand gerammten, zweifach verankerten Spundwand*, 1963
- 2 Marotz, Günter: *Beitrag zur Frage der Standfestigkeit von dichten Asphaltbelägen im Großwasserbau*, 1964
- 3 Gurr, Siegfried: *Beitrag zur Berechnung zusammengesetzter ebener Flächentragwerke unter besonderer Berücksichtigung ebener Stauwände, mit Hilfe von Randwert- und Lastwertmatrizen*, 1965
- 4 Plica, Peter: *Ein Beitrag zur Anwendung von Schalenkonstruktionen im Stahlwasserbau*, und Petrikat, Kurt: *Möglichkeiten und Grenzen des wasserbaulichen Versuchswesens*, 1966

- 5 Plate, Erich: *Beitrag zur Bestimmung der Windgeschwindigkeitsverteilung in der durch eine Wand gestörten bodennahen Luftschicht*, und
Röhnisch, Arthur; Marotz, Günter: *Neue Baustoffe und Bauausführungen für den Schutz der Böschungen und der Sohle von Kanälen, Flüssen und Häfen; Gesteungskosten und jeweilige Vorteile*, sowie
Unny, T.E.: *Schwingungsuntersuchungen am Kegelstrahlschieber*, 1967
- 6 Seiler, Erich: *Die Ermittlung des Anlagenwertes der bundeseigenen Binnenschiffahrtsstraßen und Talsperren und des Anteils der Binnenschifffahrt an diesem Wert*, 1967
- 7 *Sonderheft anlässlich des 65. Geburtstages von Prof. Arthur Röhnisch mit Beiträgen von*
Benk, Dieter; Breitling, J.; Gurr, Siegfried; Haberhauer, Robert; Honekamp, Hermann; Kuz, Klaus Dieter; Marotz, Günter; Mayer-Vorfelder, Hans-Jörg; Miller, Rudolf; Plate, Erich J.; Radomski, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1967
- 8 Jumikis, Alfred: *Beitrag zur experimentellen Untersuchung des Wassernachschubs in einem gefrierenden Boden und die Beurteilung der Ergebnisse*, 1968
- 9 Marotz, Günter: *Technische Grundlagen einer Wasserspeicherung im natürlichen Untergrund*, 1968
- 10 Radomski, Helge: *Untersuchungen über den Einfluß der Querschnittsform wellenförmiger Spundwände auf die statischen und rammtechnischen Eigenschaften*, 1968
- 11 Schwarz, Helmut: *Die Grenztragfähigkeit des Baugrundes bei Einwirkung vertikal gezogener Ankerplatten als zweidimensionales Bruchproblem*, 1969
- 12 Erbel, Klaus: *Ein Beitrag zur Untersuchung der Metamorphose von Mittelgebirgsschneedecken unter besonderer Berücksichtigung eines Verfahrens zur Bestimmung der thermischen Schneequalität*, 1969
- 13 Westhaus, Karl-Heinz: *Der Strukturwandel in der Binnenschifffahrt und sein Einfluß auf den Ausbau der Binnenschiffskanäle*, 1969
- 14 Mayer-Vorfelder, Hans-Jörg: *Ein Beitrag zur Berechnung des Erdwiderstandes unter Ansatz der logarithmischen Spirale als Gleitflächenfunktion*, 1970
- 15 Schulz, Manfred: *Berechnung des räumlichen Erddruckes auf die Wandung kreiszylindrischer Körper*, 1970
- 16 Mobasserri, Manoutschehr: *Die Rippenstützmauer. Konstruktion und Grenzen ihrer Standicherheit*, 1970
- 17 Benk, Dieter: *Ein Beitrag zum Betrieb und zur Bemessung von Hochwasserrückhaltebecken*, 1970
- 18 Gàl, Attila: *Bestimmung der mitschwingenden Wassermasse bei überströmten Fischbauchklappen mit kreiszylindrischem Staublech*, 1971, vergriffen
- 19 Kuz, Klaus Dieter: *Ein Beitrag zur Frage des Einsetzens von Kavitationserscheinungen in einer Düsenströmung bei Berücksichtigung der im Wasser gelösten Gase*, 1971, vergriffen
- 20 Schaak, Hartmut: *Verteilleitungen von Wasserkraftanlagen*, 1971
- 21 *Sonderheft zur Eröffnung der neuen Versuchsanstalt des Instituts für Wasserbau der Universität Stuttgart mit Beiträgen von* Brombach, Hansjörg; Dirksen, Wolfram; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Holthoff, Franz-Josef; Kuz, Klaus Dieter; Marotz, Günter; Minor, Hans-Erwin; Petrikat, Kurt; Röhnisch, Arthur; Rueff, Helge; Schwarz, Helmut; Vollmer, Ernst; Wildenhahn, Eberhard; 1972
- 22 Wang, Chung-su: *Ein Beitrag zur Berechnung der Schwingungen an Kegelstrahlschiebern*, 1972
- 23 Mayer-Vorfelder, Hans-Jörg: *Erdwiderstandsbeiwerte nach dem Ohde-Variationsverfahren*, 1972
- 24 Minor, Hans-Erwin: *Beitrag zur Bestimmung der Schwingungsanfachungsfunktionen überströmter Stauklappen*, 1972, vergriffen
- 25 Brombach, Hansjörg: *Untersuchung strömungsmechanischer Elemente (Fluidik) und die Möglichkeit der Anwendung von Wirbelkammerelementen im Wasserbau*, 1972, vergriffen
- 26 Wildenhahn, Eberhard: *Beitrag zur Berechnung von Horizontalfilterbrunnen*, 1972

- 27 Steinlein, Helmut: *Die Eliminierung der Schwebstoffe aus Flußwasser zum Zweck der unterirdischen Wasserspeicherung, gezeigt am Beispiel der Iller*, 1972
- 28 Holthoff, Franz Josef: *Die Überwindung großer Hubhöhen in der Binnenschifffahrt durch Schwimmerhebwerke*, 1973
- 29 Röder, Karl: *Einwirkungen aus Baugrundbewegungen auf trog- und kastenförmige Konstruktionen des Wasser- und Tunnelbaues*, 1973
- 30 Kretschmer, Heinz: *Die Bemessung von Bogenstaumauern in Abhängigkeit von der Talform*, 1973
- 31 Honekamp, Hermann: *Beitrag zur Berechnung der Montage von Unterwasserpipelines*, 1973
- 32 Giesecke, Jürgen: *Die Wirbelkammertriode als neuartiges Steuerorgan im Wasserbau*, und Brombach, Hansjörg: *Entwicklung, Bauformen, Wirkungsweise und Steuereigenschaften von Wirbelkammerverstärkern*, 1974
- 33 Rueff, Helge: *Untersuchung der schwingungserregenden Kräfte an zwei hintereinander angeordneten Tiefschützen unter besonderer Berücksichtigung von Kavitation*, 1974
- 34 Röhnisch, Arthur: *Einpreßversuche mit Zementmörtel für Spannbeton - Vergleich der Ergebnisse von Modellversuchen mit Ausführungen in Hüllwellrohren*, 1975
- 35 *Sonderheft anlässlich des 65. Geburtstages von Prof. Dr.-Ing. Kurt Petrikat mit Beiträgen von:* Brombach, Hansjörg; Erbel, Klaus; Flinspach, Dieter; Fischer jr., Richard; Gàl, Attila; Gerlach, Reinhard; Giesecke, Jürgen; Haberhauer, Robert; Hafner Edzard; Hausenblas, Bernhard; Horlacher, Hans-Burkhard; Hutarew, Andreas; Knoll, Manfred; Krummet, Ralph; Marotz, Günter; Merkle, Theodor; Miller, Christoph; Minor, Hans-Erwin; Neumayer, Hans; Rao, Syamala; Rath, Paul; Rueff, Helge; Ruppert, Jürgen; Schwarz, Wolfgang; Topal-Gökceli, Mehmet; Vollmer, Ernst; Wang, Chung-su; Weber, Hans-Georg; 1975
- 36 Berger, Jochum: *Beitrag zur Berechnung des Spannungszustandes in rotationssymmetrisch belasteten Kugelschalen veränderlicher Wandstärke unter Gas- und Flüssigkeitsdruck durch Integration schwach singulärer Differentialgleichungen*, 1975
- 37 Dirksen, Wolfram: *Berechnung instationärer Abflußvorgänge in gestauten Gerinnen mittels Differenzenverfahren und die Anwendung auf Hochwasserrückhaltebecken*, 1976
- 38 Horlacher, Hans-Burkhard: *Berechnung instationärer Temperatur- und Wärmespannungsfelder in langen mehrschichtigen Hohlzylindern*, 1976
- 39 Hafner, Edzard: *Untersuchung der hydrodynamischen Kräfte auf Baukörper im Tiefwasserbereich des Meeres*, 1977, ISBN 3-921694-39-6
- 40 Ruppert, Jürgen: *Über den Axialwirbelkammerverstärker für den Einsatz im Wasserbau*, 1977, ISBN 3-921694-40-X
- 41 Hutarew, Andreas: *Beitrag zur Beeinflussbarkeit des Sauerstoffgehalts in Fließgewässern an Abstürzen und Wehren*, 1977, ISBN 3-921694-41-8, vergriffen
- 42 Miller, Christoph: *Ein Beitrag zur Bestimmung der schwingungserregenden Kräfte an unterströmten Wehren*, 1977, ISBN 3-921694-42-6
- 43 Schwarz, Wolfgang: *Druckstoßberechnung unter Berücksichtigung der Radial- und Längsverschiebungen der Rohrwandung*, 1978, ISBN 3-921694-43-4
- 44 Kinzelbach, Wolfgang: *Numerische Untersuchungen über den optimalen Einsatz variabler Kühlsysteme einer Kraftwerkskette am Beispiel Oberrhein*, 1978, ISBN 3-921694-44-2
- 45 Barczewski, Baldur: *Neue Meßmethoden für Wasser-Luftgemische und deren Anwendung auf zweiphasige Auftriebsstrahlen*, 1979, ISBN 3-921694-45-0
- 46 Neumayer, Hans: *Untersuchung der Strömungsvorgänge in radialen Wirbelkammerverstärkern*, 1979, ISBN 3-921694-46-9
- 47 Elalfy, Youssef-Elhassan: *Untersuchung der Strömungsvorgänge in Wirbelkammerdioden und -drosseln*, 1979, ISBN 3-921694-47-7
- 48 Brombach, Hansjörg: *Automatisierung der Bewirtschaftung von Wasserspeichern*, 1981, ISBN 3-921694-48-5
- 49 Geldner, Peter: *Deterministische und stochastische Methoden zur Bestimmung der Selbstdichtung von Gewässern*, 1981, ISBN 3-921694-49-3, vergriffen

- 50 Mehlhorn, Hans: *Temperaturveränderungen im Grundwasser durch Brauchwassereinleitungen*, 1982, ISBN 3-921694-50-7, vergriffen
- 51 Hafner, Edzard: *Rohrleitungen und Behälter im Meer*, 1983, ISBN 3-921694-51-5
- 52 Rinnert, Bernd: *Hydrodynamische Dispersion in porösen Medien: Einfluß von Dichteunterschieden auf die Vertikalvermischung in horizontaler Strömung*, 1983, ISBN 3-921694-52-3, vergriffen
- 53 Lindner, Wulf: *Steuerung von Grundwasserentnahmen unter Einhaltung ökologischer Kriterien*, 1983, ISBN 3-921694-53-1, vergriffen
- 54 Herr, Michael; Herzer, Jörg; Kinzelbach, Wolfgang; Kobus, Helmut; Rinnert, Bernd: *Methoden zur rechnerischen Erfassung und hydraulischen Sanierung von Grundwasserkontaminationen*, 1983, ISBN 3-921694-54-X
- 55 Schmitt, Paul: *Wege zur Automatisierung der Niederschlagsermittlung*, 1984, ISBN 3-921694-55-8, vergriffen
- 56 Müller, Peter: *Transport und selektive Sedimentation von Schwebstoffen bei gestautem Abfluß*, 1985, ISBN 3-921694-56-6
- 57 El-Qawasmeh, Fuad: *Möglichkeiten und Grenzen der Tropfbewässerung unter besonderer Berücksichtigung der Verstopfungsanfälligkeit der Tropfelemente*, 1985, ISBN 3-921694-57-4, vergriffen
- 58 Kirchenbaur, Klaus: *Mikroprozessorgesteuerte Erfassung instationärer Druckfelder am Beispiel seegangsbelasteter Baukörper*, 1985, ISBN 3-921694-58-2
- 59 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1984/85 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart), 1985, ISBN 3-921694-59-0, vergriffen
- 60 Spitz, Karlheinz: *Dispersion in porösen Medien: Einfluß von Inhomogenitäten und Dichteunterschieden*, 1985, ISBN 3-921694-60-4, vergriffen
- 61 Kobus, Helmut: *An Introduction to Air-Water Flows in Hydraulics*, 1985, ISBN 3-921694-61-2
- 62 Kaleris, Vassilios: *Erfassung des Austausches von Oberflächen- und Grundwasser in horizontalebene Grundwassermodellen*, 1986, ISBN 3-921694-62-0
- 63 Herr, Michael: *Grundlagen der hydraulischen Sanierung verunreinigter Porengrundwasserleiter*, 1987, ISBN 3-921694-63-9
- 64 Marx, Walter: *Berechnung von Temperatur und Spannung in Massenbeton infolge Hydratation*, 1987, ISBN 3-921694-64-7
- 65 Koschitzky, Hans-Peter: *Dimensionierungskonzept für Sohlbelüfter in Schußrinnen zur Vermeidung von Kavitationsschäden*, 1987, ISBN 3-921694-65-5
- 66 Kobus, Helmut (Hrsg.): *Modellierung des großräumigen Wärme- und Schadstofftransports im Grundwasser*, Tätigkeitsbericht 1986/87 (DFG-Forschergruppe an den Universitäten Hohenheim, Karlsruhe und Stuttgart) 1987, ISBN 3-921694-66-3
- 67 Söll, Thomas: *Berechnungsverfahren zur Abschätzung anthropogener Temperaturanomalien im Grundwasser*, 1988, ISBN 3-921694-67-1
- 68 Dittrich, Andreas; Westrich, Bernd: *Bodenseeufenerosion, Bestandsaufnahme und Bewertung*, 1988, ISBN 3-921694-68-X, vergriffen
- 69 Huwe, Bernd; van der Ploeg, Rienk R.: *Modelle zur Simulation des Stickstoffhaushaltes von Standorten mit unterschiedlicher landwirtschaftlicher Nutzung*, 1988, ISBN 3-921694-69-8, vergriffen
- 70 Stephan, Karl: *Integration elliptischer Funktionen*, 1988, ISBN 3-921694-70-1
- 71 Kobus, Helmut; Zilliox, Lothaire (Hrsg.): *Nitratbelastung des Grundwassers, Auswirkungen der Landwirtschaft auf die Grundwasser- und Rohwasserbeschaffenheit und Maßnahmen zum Schutz des Grundwassers*. Vorträge des deutsch-französischen Kolloquiums am 6. Oktober 1988, Universitäten Stuttgart und Louis Pasteur Strasbourg (Vorträge in deutsch oder französisch, Kurzfassungen zweisprachig), 1988, ISBN 3-921694-71-X

- 72 Soyeaux, Renald: *Unterströmung von Stauanlagen auf klüftigem Untergrund unter Berücksichtigung laminarer und turbulenter Fließzustände*, 1991, ISBN 3-921694-72-8
- 73 Kohane, Roberto: *Berechnungsmethoden für Hochwasserabfluß in Fließgewässern mit überströmten Vorländern*, 1991, ISBN 3-921694-73-6
- 74 Hassinger, Reinhard: *Beitrag zur Hydraulik und Bemessung von Blocksteinrampen in flexibler Bauweise*, 1991, ISBN 3-921694-74-4, vergriffen
- 75 Schäfer, Gerhard: *Einfluß von Schichtenstrukturen und lokalen Einlagerungen auf die Längsdispersion in Porengrundwasserleitern*, 1991, ISBN 3-921694-75-2
- 76 Giesecke, Jürgen: *Vorträge, Wasserwirtschaft in stark besiedelten Regionen; Umweltforschung mit Schwerpunkt Wasserwirtschaft*, 1991, ISBN 3-921694-76-0
- 77 Huwe, Bernd: *Deterministische und stochastische Ansätze zur Modellierung des Stickstoffhaushalts landwirtschaftlich genutzter Flächen auf unterschiedlichem Skalenniveau*, 1992, ISBN 3-921694-77-9, vergriffen
- 78 Rommel, Michael: *Verwendung von Kluffdaten zur realitätsnahen Generierung von Kluffnetzen mit anschließender laminar-turbulenter Strömungsberechnung*, 1993, ISBN 3-92 1694-78-7
- 79 Marschall, Paul: *Die Ermittlung lokaler Stofffrachten im Grundwasser mit Hilfe von Einbohrloch-Meßverfahren*, 1993, ISBN 3-921694-79-5, vergriffen
- 80 Ptak, Thomas: *Stofftransport in heterogenen Porenaquiferen: Felduntersuchungen und stochastische Modellierung*, 1993, ISBN 3-921694-80-9, vergriffen
- 81 Haakh, Frieder: *Transientes Strömungsverhalten in Wirbelkammern*, 1993, ISBN 3-921694-81-7
- 82 Kobus, Helmut; Cirpka, Olaf; Barczewski, Baldur; Koschitzky, Hans-Peter: *Versuchseinrichtung zur Grundwasser und Altlastensanierung VEGAS, Konzeption und Programmrahmen*, 1993, ISBN 3-921694-82-5
- 83 Zang, Weidong: *Optimaler Echtzeit-Betrieb eines Speichers mit aktueller Abflußregenerierung*, 1994, ISBN 3-921694-83-3, vergriffen
- 84 Franke, Hans-Jörg: *Stochastische Modellierung eines flächenhaften Stoffeintrages und Transports in Grundwasser am Beispiel der Pflanzenschutzmittelproblematik*, 1995, ISBN 3-921694-84-1
- 85 Lang, Ulrich: *Simulation regionaler Strömungs- und Transportvorgänge in Karstaquiferen mit Hilfe des Doppelkontinuum-Ansatzes: Methodenentwicklung und Parameteridentifikation*, 1995, ISBN 3-921694-85-X, vergriffen
- 86 Helmig, Rainer: *Einführung in die Numerischen Methoden der Hydromechanik*, 1996, ISBN 3-921694-86-8, vergriffen
- 87 Cirpka, Olaf: *CONTRACT: A Numerical Tool for Contaminant Transport and Chemical Transformations - Theory and Program Documentation -*, 1996, ISBN 3-921694-87-6
- 88 Haberlandt, Uwe: *Stochastische Synthese und Regionalisierung des Niederschlages für Schmutzfrachtberechnungen*, 1996, ISBN 3-921694-88-4
- 89 Croisé, Jean: *Extraktion von flüchtigen Chemikalien aus natürlichen Lockergesteinen mittels erzwungener Luftströmung*, 1996, ISBN 3-921694-89-2, vergriffen
- 90 Jorde, Klaus: *Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken*, 1997, ISBN 3-921694-90-6, vergriffen
- 91 Helmig, Rainer: *Gekoppelte Strömungs- und Transportprozesse im Untergrund - Ein Beitrag zur Hydrosystemmodellierung-*, 1998, ISBN 3-921694-91-4, vergriffen
- 92 Emmert, Martin: *Numerische Modellierung nichtisothermer Gas-Wasser Systeme in porösen Medien*, 1997, ISBN 3-921694-92-2
- 93 Kern, Ulrich: *Transport von Schweb- und Schadstoffen in staugeregelten Fließgewässern am Beispiel des Neckars*, 1997, ISBN 3-921694-93-0, vergriffen
- 94 Förster, Georg: *Druckstoßdämpfung durch große Luftblasen in Hochpunkten von Rohrleitungen* 1997, ISBN 3-921694-94-9

- 95 Cirpka, Olaf: *Numerische Methoden zur Simulation des reaktiven Mehrkomponententransports im Grundwasser*, 1997, ISBN 3-921694-95-7, vergriffen
- 96 Färber, Arne: *Wärmetransport in der ungesättigten Bodenzone: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1997, ISBN 3-921694-96-5
- 97 Betz, Christoph: *Wasserdampfdestillation von Schadstoffen im porösen Medium: Entwicklung einer thermischen In-situ-Sanierungstechnologie*, 1998, SBN 3-921694-97-3
- 98 Xu, Yichun: *Numerical Modeling of Suspended Sediment Transport in Rivers*, 1998, ISBN 3-921694-98-1, vergriffen
- 99 Wüst, Wolfgang: *Geochemische Untersuchungen zur Sanierung CKW-kontaminierter Aquifere mit Fe(0)-Reaktionswänden*, 2000, ISBN 3-933761-02-2
- 100 Sheta, Hussam: *Simulation von Mehrphasenvorgängen in porösen Medien unter Einbeziehung von Hysterese-Effekten*, 2000, ISBN 3-933761-03-4
- 101 Ayros, Edwin: *Regionalisierung extremer Abflüsse auf der Grundlage statistischer Verfahren*, 2000, ISBN 3-933761-04-2, vergriffen
- 102 Huber, Ralf: *Compositional Multiphase Flow and Transport in Heterogeneous Porous Media*, 2000, ISBN 3-933761-05-0
- 103 Braun, Christopherus: *Ein Upscaling-Verfahren für Mehrphasenströmungen in porösen Medien*, 2000, ISBN 3-933761-06-9
- 104 Hofmann, Bernd: *Entwicklung eines rechnergestützten Managementsystems zur Beurteilung von Grundwasserschadensfällen*, 2000, ISBN 3-933761-07-7
- 105 Class, Holger: *Theorie und numerische Modellierung nichtisothermer Mehrphasenprozesse in NAPL-kontaminierten porösen Medien*, 2001, ISBN 3-933761-08-5
- 106 Schmidt, Reinhard: *Wasserdampf- und Heißluftinjektion zur thermischen Sanierung kontaminierter Standorte*, 2001, ISBN 3-933761-09-3
- 107 Josef, Reinhold: *Schadstoffextraktion mit hydraulischen Sanierungsverfahren unter Anwendung von grenzflächenaktiven Stoffen*, 2001, ISBN 3-933761-10-7
- 108 Schneider, Matthias: *Habitat- und Abflussmodellierung für Fließgewässer mit unscharfen Berechnungsansätzen*, 2001, ISBN 3-933761-11-5
- 109 Rathgeb, Andreas: *Hydrodynamische Bemessungsgrundlagen für Lockerdeckwerke an überströmbaren Erdämmen*, 2001, ISBN 3-933761-12-3
- 110 Lang, Stefan: *Parallele numerische Simulation instationärer Probleme mit adaptiven Methoden auf unstrukturierten Gittern*, 2001, ISBN 3-933761-13-1
- 111 Appt, Jochen; Stumpp Simone: *Die Bodensee-Messkampagne 2001, IWS/CWR Lake Constance Measurement Program 2001*, 2002, ISBN 3-933761-14-X
- 112 Heimerl, Stephan: *Systematische Beurteilung von Wasserkraftprojekten*, 2002, ISBN 3-933761-15-8, vergriffen
- 113 Iqbal, Amin: *On the Management and Salinity Control of Drip Irrigation*, 2002, ISBN 3-933761-16-6
- 114 Silberhorn-Hemminger, Annette: *Modellierung von Klufftaquifersystemen: Geostatistische Analyse und deterministisch-stochastische Kluffgenerierung*, 2002, ISBN 3-933761-17-4
- 115 Winkler, Angela: *Prozesse des Wärme- und Stofftransports bei der In-situ-Sanierung mit festen Wärmequellen*, 2003, ISBN 3-933761-18-2
- 116 Marx, Walter: *Wasserkraft, Bewässerung, Umwelt - Planungs- und Bewertungsschwerpunkte der Wasserbewirtschaftung*, 2003, ISBN 3-933761-19-0
- 117 Hinkelmann, Reinhard: *Efficient Numerical Methods and Information-Processing Techniques in Environment Water*, 2003, ISBN 3-933761-20-4
- 118 Samaniego-Eguiguren, Luis Eduardo: *Hydrological Consequences of Land Use / Land Cover and Climatic Changes in Mesoscale Catchments*, 2003, ISBN 3-933761-21-2
- 119 Neunhäuserer, Lina: *Diskretisierungsansätze zur Modellierung von Strömungs- und Transportprozessen in geklüftet-porösen Medien*, 2003, ISBN 3-933761-22-0
- 120 Paul, Maren: *Simulation of Two-Phase Flow in Heterogeneous Poros Media with Adaptive Methods*, 2003, ISBN 3-933761-23-9

- 121 Ehret, Uwe: *Rainfall and Flood Nowcasting in Small Catchments using Weather Radar*, 2003, ISBN 3-933761-24-7
- 122 Haag, Ingo: *Der Sauerstoffhaushalt staugeregelter Flüsse am Beispiel des Neckars - Analysen, Experimente, Simulationen -*, 2003, ISBN 3-933761-25-5
- 123 Appt, Jochen: *Analysis of Basin-Scale Internal Waves in Upper Lake Constance*, 2003, ISBN 3-933761-26-3
- 124 Hrsg.: Schrenk, Volker; Batereau, Katrin; Barczewski, Baldur; Weber, Karolin und Koschitzky, Hans-Peter: *Symposium Ressource Fläche und VEGAS - Statuskolloquium 2003, 30. September und 1. Oktober 2003*, 2003, ISBN 3-933761-27-1
- 125 Omar Khalil Ouda: *Optimisation of Agricultural Water Use: A Decision Support System for the Gaza Strip*, 2003, ISBN 3-933761-28-0
- 126 Batereau, Katrin: *Sensorbasierte Bodenluftmessung zur Vor-Ort-Erkundung von Schadensherden im Untergrund*, 2004, ISBN 3-933761-29-8
- 127 Witt, Oliver: *Erosionsstabilität von Gewässersedimenten mit Auswirkung auf den Stofftransport bei Hochwasser am Beispiel ausgewählter Stauhaltungen des Oberrheins*, 2004, ISBN 3-933761-30-1
- 128 Jakobs, Hartmut: *Simulation nicht-isothermer Gas-Wasser-Prozesse in komplexen Kluft-Matrix-Systemen*, 2004, ISBN 3-933761-31-X
- 129 Li, Chen-Chien: *Deterministisch-stochastisches Berechnungskonzept zur Beurteilung der Auswirkungen erosiver Hochwasserereignisse in Flusstauhaltungen*, 2004, ISBN 3-933761-32-8
- 130 Reichenberger, Volker; Helmig, Rainer; Jakobs, Hartmut; Bastian, Peter; Niessner, Jennifer: *Complex Gas-Water Processes in Discrete Fracture-Matrix Systems: Up-scaling, Mass-Conservative Discretization and Efficient Multilevel Solution*, 2004, ISBN 3-933761-33-6
- 131 Hrsg.: Barczewski, Baldur; Koschitzky, Hans-Peter; Weber, Karolin; Wege, Ralf: *VEGAS - Statuskolloquium 2004*, Tagungsband zur Veranstaltung am 05. Oktober 2004 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2004, ISBN 3-933761-34-4
- 132 Asie, Kemal Jabir: *Finite Volume Models for Multiphase Multicomponent Flow through Porous Media*. 2005, ISBN 3-933761-35-2
- 133 Jacoub, George: *Development of a 2-D Numerical Module for Particulate Contaminant Transport in Flood Retention Reservoirs and Impounded Rivers*, 2004, ISBN 3-933761-36-0
- 134 Nowak, Wolfgang: *Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface*, 2005, ISBN 3-933761-37-9
- 135 Süß, Mia: *Analysis of the influence of structures and boundaries on flow and transport processes in fractured porous media*, 2005, ISBN 3-933761-38-7
- 136 Jose, Surabhin Chackiath: *Experimental Investigations on Longitudinal Dispersive Mixing in Heterogeneous Aquifers*, 2005, ISBN: 3-933761-39-5
- 137 Filiz, Fulya: *Linking Large-Scale Meteorological Conditions to Floods in Mesoscale Catchments*, 2005, ISBN 3-933761-40-9
- 138 Qin, Minghao: *Wirklichkeitsnahe und recheneffiziente Ermittlung von Temperatur und Spannungen bei großen RCC-Staumauern*, 2005, ISBN 3-933761-41-7
- 139 Kobayashi, Kenichiro: *Optimization Methods for Multiphase Systems in the Subsurface - Application to Methane Migration in Coal Mining Areas*, 2005, ISBN 3-933761-42-5
- 140 Rahman, Md. Arifur: *Experimental Investigations on Transverse Dispersive Mixing in Heterogeneous Porous Media*, 2005, ISBN 3-933761-43-3
- 141 Schrenk, Volker: *Ökobilanzen zur Bewertung von Altlastensanierungsmaßnahmen*, 2005, ISBN 3-933761-44-1
- 142 Hundecha, Hirpa Yeshewatesfa: *Regionalization of Parameters of a Conceptual Rainfall-Runoff Model*, 2005, ISBN: 3-933761-45-X
- 143 Wege, Ralf: *Untersuchungs- und Überwachungsmethoden für die Beurteilung natürlicher Selbstreinigungsprozesse im Grundwasser*, 2005, ISBN 3-933761-46-8

- 144 Breiting, Thomas: *Techniken und Methoden der Hydroinformatik - Modellierung von komplexen Hydrosystemen im Untergrund*, 2006, ISBN 3-933761-47-6
- 145 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Müller, Martin: *Ressource Untergrund: 10 Jahre VEGAS: Forschung und Technologieentwicklung zum Schutz von Grundwasser und Boden*, Tagungsband zur Veranstaltung am 28. und 29. September 2005 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2005, ISBN 3-933761-48-4
- 146 Rojanschi, Vlad: *Abflusskonzentration in mesoskaligen Einzugsgebieten unter Berücksichtigung des Sickerraumes*, 2006, ISBN 3-933761-49-2
- 147 Winkler, Nina Simone: *Optimierung der Steuerung von Hochwasserrückhaltebecken-systemen*, 2006, ISBN 3-933761-50-6
- 148 Wolf, Jens: *Räumlich differenzierte Modellierung der Grundwasserströmung alluvialer Aquifere für mesoskalige Einzugsgebiete*, 2006, ISBN: 3-933761-51-4
- 149 Kohler, Beate: *Externe Effekte der Laufwasserkraftnutzung*, 2006, ISBN 3-933761-52-2
- 150 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhmann, Matthias: *VEGAS-Statuskolloquium 2006*, Tagungsband zur Veranstaltung am 28. September 2006 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2006, ISBN 3-933761-53-0
- 151 Niessner, Jennifer: *Multi-Scale Modeling of Multi-Phase - Multi-Component Processes in Heterogeneous Porous Media*, 2006, ISBN 3-933761-54-9
- 152 Fischer, Markus: *Beanspruchung eingeeerdeter Rohrleitungen infolge Austrocknung bindiger Böden*, 2006, ISBN 3-933761-55-7
- 153 Schneck, Alexander: *Optimierung der Grundwasserbewirtschaftung unter Berücksichtigung der Belange der Wasserversorgung, der Landwirtschaft und des Naturschutzes*, 2006, ISBN 3-933761-56-5
- 154 Das, Tapash: *The Impact of Spatial Variability of Precipitation on the Predictive Uncertainty of Hydrological Models*, 2006, ISBN 3-33761-57-3
- 155 Bielinski, Andreas: *Numerical Simulation of CO₂ sequestration in geological formations*, 2007, ISBN 3-933761-58-1
- 156 Mödinger, Jens: *Entwicklung eines Bewertungs- und Entscheidungsunterstützungssysteme für eine nachhaltige regionale Grundwasserbewirtschaftung*, 2006, ISBN 3-933761-60-3
- 157 Manthey, Sabine: *Two-phase flow processes with dynamic effects in porous media - parameter estimation and simulation*, 2007, ISBN 3-933761-61-1
- 158 Pozos Estrada, Oscar: *Investigation on the Effects of Entrained Air in Pipelines*, 2007, ISBN 3-933761-62-X
- 159 Ochs, Steffen Oliver: *Steam injection into saturated porous media – process analysis including experimental and numerical investigations*, 2007, ISBN 3-933761-63-8
- 160 Marx, Andreas: *Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage*, 2007, ISBN 3-933761-64-6
- 161 Hartmann, Gabriele Maria: *Investigation of Evapotranspiration Concepts in Hydrological Modelling for Climate Change Impact Assessment*, 2007, ISBN 3-933761-65-4
- 162 Kebede Gurmessa, Tesfaye: *Numerical Investigation on Flow and Transport Characteristics to Improve Long-Term Simulation of Reservoir Sedimentation*, 2007, ISBN 3-933761-66-2
- 163 Trifković, Aleksandar: *Multi-objective and Risk-based Modelling Methodology for Planning, Design and Operation of Water Supply Systems*, 2007, ISBN 3-933761-67-0
- 164 Götzinger, Jens: *Distributed Conceptual Hydrological Modelling - Simulation of Climate, Land Use Change Impact and Uncertainty Analysis*, 2007, ISBN 3-933761-68-9
- 165 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhmann, Matthias: *VEGAS – Kolloquium 2007*, Tagungsband zur Veranstaltung am 26. September 2007 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2007, ISBN 3-933761-69-7
- 166 Freeman, Beau: *Modernization Criteria Assessment for Water Resources Planning; Klamath Irrigation Project, U.S.*, 2008, ISBN 3-933761-70-0

- 167 Dreher, Thomas: *Selektive Sedimentation von Feinstschwebstoffen in Wechselwirkung mit wandnahen turbulenten Strömungsbedingungen*, 2008, ISBN 3-933761-71-9
- 168 Yang, Wei: *Discrete-Continuous Downscaling Model for Generating Daily Precipitation Time Series*, 2008, ISBN 3-933761-72-7
- 169 Kopecki, Ianina: *Calculational Approach to FST-Hemispheres for Multiparametrical Benthos Habitat Modelling*, 2008, ISBN 3-933761-73-5
- 170 Brommundt, Jürgen: *Stochastische Generierung räumlich zusammenhängender Niederschlagszeitreihen*, 2008, ISBN 3-933761-74-3
- 171 Papafotiou, Alexandros: *Numerical Investigations of the Role of Hysteresis in Heterogeneous Two-Phase Flow Systems*, 2008, ISBN 3-933761-75-1
- 172 He, Yi: *Application of a Non-Parametric Classification Scheme to Catchment Hydrology*, 2008, ISBN 978-3-933761-76-7
- 173 Wagner, Sven: *Water Balance in a Poorly Gauged Basin in West Africa Using Atmospheric Modelling and Remote Sensing Information*, 2008, ISBN 978-3-933761-77-4
- 174 Hrsg.: Braun, Jürgen; Koschitzky, Hans-Peter; Stuhmann, Matthias; Schrenk, Volker: *VEGAS-Kolloquium 2008 Ressource Fläche III*, Tagungsband zur Veranstaltung am 01. Oktober 2008 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2008, ISBN 978-3-933761-78-1
- 175 Patil, Sachin: *Regionalization of an Event Based Nash Cascade Model for Flood Predictions in Ungauged Basins*, 2008, ISBN 978-3-933761-79-8
- 176 Assteerawatt, Anongnart: *Flow and Transport Modelling of Fractured Aquifers based on a Geostatistical Approach*, 2008, ISBN 978-3-933761-80-4
- 177 Karnahl, Joachim Alexander: *2D numerische Modellierung von multifraktionalem Schwebstoff- und Schadstofftransport in Flüssen*, 2008, ISBN 978-3-933761-81-1
- 178 Hiester, Uwe: *Technologieentwicklung zur In-situ-Sanierung der ungesättigten Bodenzone mit festen Wärmequellen*, 2009, ISBN 978-3-933761-82-8
- 179 Laux, Patrick: *Statistical Modeling of Precipitation for Agricultural Planning in the Volta Basin of West Africa*, 2009, ISBN 978-3-933761-83-5
- 180 Ehsan, Saqib: *Evaluation of Life Safety Risks Related to Severe Flooding*, 2009, ISBN 978-3-933761-84-2
- 181 Prohaska, Sandra: *Development and Application of a 1D Multi-Strip Fine Sediment Transport Model for Regulated Rivers*, 2009, ISBN 978-3-933761-85-9
- 182 Kopp, Andreas: *Evaluation of CO₂ Injection Processes in Geological Formations for Site Screening*, 2009, ISBN 978-3-933761-86-6
- 183 Ebigbo, Anozie: *Modelling of biofilm growth and its influence on CO₂ and water (two-phase) flow in porous media*, 2009, ISBN 978-3-933761-87-3
- 184 Freiboth, Sandra: *A phenomenological model for the numerical simulation of multiphase multicomponent processes considering structural alterations of porous media*, 2009, ISBN 978-3-933761-88-0
- 185 Zöllner, Frank: *Implementierung und Anwendung netzfreier Methoden im Konstruktiven Wasserbau und in der Hydromechanik*, 2009, ISBN 978-3-933761-89-7
- 186 Vasin, Milos: *Influence of the soil structure and property contrast on flow and transport in the unsaturated zone*, 2010, ISBN 978-3-933761-90-3
- 187 Li, Jing: *Application of Copulas as a New Geostatistical Tool*, 2010, ISBN 978-3-933761-91-0
- 188 AghaKouchak, Amir: *Simulation of Remotely Sensed Rainfall Fields Using Copulas*, 2010, ISBN 978-3-933761-92-7
- 189 Thapa, Pawan Kumar: *Physically-based spatially distributed rainfall runoff modelling for soil erosion estimation*, 2010, ISBN 978-3-933761-93-4
- 190 Wurms, Sven: *Numerische Modellierung der Sedimentationsprozesse in Retentionsanlagen zur Steuerung von Stoffströmen bei extremen Hochwasserabflussereignissen*, 2011, ISBN 978-3-933761-94-1

- 191 Merkel, Uwe: *Unsicherheitsanalyse hydraulischer Einwirkungen auf Hochwasserschutzdeiche und Steigerung der Leistungsfähigkeit durch adaptive Strömungsmodellierung*, 2011, ISBN 978-3-933761-95-8
- 192 Fritz, Jochen: *A Decoupled Model for Compositional Non-Isothermal Multiphase Flow in Porous Media and Multiphysics Approaches for Two-Phase Flow*, 2010, ISBN 978-3-933761-96-5
- 193 Weber, Karolin (Hrsg.): *12. Treffen junger WissenschaftlerInnen an Wasserbauinstituten*, 2010, ISBN 978-3-933761-97-2
- 194 Bliedernicht, Jan-Geert: *Probability Forecasts of Daily Areal Precipitation for Small River Basins*, 2011, ISBN 978-3-933761-98-9
- 195 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2010 In-situ-Sanierung - Stand und Entwicklung Nano und ISCO -*, Tagungsband zur Veranstaltung am 07. Oktober 2010 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2010, ISBN 978-3-933761-99-6
- 196 Gafurov, Abror: *Water Balance Modeling Using Remote Sensing Information - Focus on Central Asia*, 2010, ISBN 978-3-942036-00-9
- 197 Mackenberg, Sylvia: *Die Quellstärke in der Sickerwasserprognose: Möglichkeiten und Grenzen von Labor- und Freilanduntersuchungen*, 2010, ISBN 978-3-942036-01-6
- 198 Singh, Shailesh Kumar: *Robust Parameter Estimation in Gauged and Ungauged Basins*, 2010, ISBN 978-3-942036-02-3
- 199 Doğan, Mehmet Onur: *Coupling of porous media flow with pipe flow*, 2011, ISBN 978-3-942036-03-0
- 200 Liu, Min: *Study of Topographic Effects on Hydrological Patterns and the Implication on Hydrological Modeling and Data Interpolation*, 2011, ISBN 978-3-942036-04-7
- 201 Geleta, Habtamu Itefa: *Watershed Sediment Yield Modeling for Data Scarce Areas*, 2011, ISBN 978-3-942036-05-4
- 202 Franke, Jörg: *Einfluss der Überwachung auf die Versagenswahrscheinlichkeit von Staustufen*, 2011, ISBN 978-3-942036-06-1
- 203 Bakimchandra, Oinam: *Integrated Fuzzy-GIS approach for assessing regional soil erosion risks*, 2011, ISBN 978-3-942036-07-8
- 204 Alam, Muhammad Mahboob: *Statistical Downscaling of Extremes of Precipitation in Mesoscale Catchments from Different RCMs and Their Effects on Local Hydrology*, 2011, ISBN 978-3-942036-08-5
- 205 Hrsg.: Koschitzky, Hans-Peter; Braun, Jürgen: *VEGAS-Kolloquium 2011 Flache Geothermie - Perspektiven und Risiken*, Tagungsband zur Veranstaltung am 06. Oktober 2011 an der Universität Stuttgart, Campus Stuttgart-Vaihingen, 2011, ISBN 978-3-933761-09-2
- 206 Haslauer, Claus: *Analysis of Real-World Spatial Dependence of Subsurface Hydraulic Properties Using Copulas with a Focus on Solute Transport Behaviour*, 2011, ISBN 978-3-942036-10-8
- 207 Dung, Nguyen Viet: *Multi-objective automatic calibration of hydrodynamic models – development of the concept and an application in the Mekong Delta*, 2011, ISBN 978-3-942036-11-5
- 208 Hung, Nguyen Nghia: *Sediment dynamics in the floodplain of the Mekong Delta, Vietnam*, 2011, ISBN 978-3-942036-12-2
- 209 Kuhlmann, Anna: *Influence of soil structure and root water uptake on flow in the unsaturated zone*, 2012, ISBN 978-3-942036-13-9
- 210 Tuhtan, Jeffrey Andrew: *Including the Second Law Inequality in Aquatic Ecodynamics: A Modeling Approach for Alpine Rivers Impacted by Hydropeaking*, 2012, ISBN 978-3-942036-14-6
- 211 Tolossa, Habtamu: *Sediment Transport Computation Using a Data-Driven Adaptive Neuro-Fuzzy Modelling Approach*, 2012, ISBN 978-3-942036-15-3

- 212 Tatomir, Alexandru-Bodgan: *From Discrete to Continuum Concepts of Flow in Fractured Porous Media*, 2012, ISBN 978-3-942036-16-0
- 213 Erbertseder, Karin: *A Multi-Scale Model for Describing Cancer-Therapeutic Transport in the Human Lung*, 2012, ISBN 978-3-942036-17-7
- 214 Noack, Markus: *Modelling Approach for Interstitial Sediment Dynamics and Reproduction of Gravel Spawning Fish*, 2012, ISBN 978-3-942036-18-4
- 215 De Boer, Cjestmir Volkert: *Transport of Nano Sized Zero Valent Iron Colloids during Injection into the Subsurface*, 2012, ISBN 978-3-942036-19-1
- 216 Pfaff, Thomas: *Processing and Analysis of Weather Radar Data for Use in Hydrology*, 2013, ISBN 978-3-942036-20-7
- 217 Lebreuz, Hans-Henning: *Addressing the Input Uncertainty for Hydrological Modeling by a New Geostatistical Method*, 2013, ISBN 978-3-942036-21-4
- 218 Darcis, Melanie Yvonne: *Coupling Models of Different Complexity for the Simulation of CO₂ Storage in Deep Saline Aquifers*, 2013, ISBN 978-3-942036-22-1
- 219 Beck, Ferdinand: *Generation of Spatially Correlated Synthetic Rainfall Time Series in High Temporal Resolution - A Data Driven Approach*, 2013, ISBN 978-3-942036-23-8
- 220 Guthke, Philipp: *Non-multi-Gaussian spatial structures: Process-driven natural genesis, manifestation, modeling approaches, and influences on dependent processes*, 2013, ISBN 978-3-942036-24-5
- 221 Walter, Lena: *Uncertainty studies and risk assessment for CO₂ storage in geological formations*, 2013, ISBN 978-3-942036-25-2
- 222 Wolff, Markus: *Multi-scale modeling of two-phase flow in porous media including capillary pressure effects*, 2013, ISBN 978-3-942036-26-9
- 223 Mosthaf, Klaus Roland: *Modeling and analysis of coupled porous-medium and free flow with application to evaporation processes*, 2014, ISBN 978-3-942036-27-6
- 224 Leube, Philipp Christoph: *Methods for Physically-Based Model Reduction in Time: Analysis, Comparison of Methods and Application*, 2013, ISBN 978-3-942036-28-3
- 225 Rodríguez Fernández, Jhan Ignacio: *High Order Interactions among environmental variables: Diagnostics and initial steps towards modeling*, 2013, ISBN 978-3-942036-29-0
- 226 Eder, Maria Magdalena: *Climate Sensitivity of a Large Lake*, 2013, ISBN 978-3-942036-30-6
- 227 Greiner, Philipp: *Alkoholinjektion zur In-situ-Sanierung von CKW Schadensherden in Grundwasserleitern: Charakterisierung der relevanten Prozesse auf unterschiedlichen Skalen*, 2014, ISBN 978-3-942036-31-3
- 228 Lauser, Andreas: *Theory and Numerical Applications of Compositional Multi-Phase Flow in Porous Media*, 2014, ISBN 978-3-942036-32-0
- 229 Enzenhöfer, Rainer: *Risk Quantification and Management in Water Production and Supply Systems*, 2014, ISBN 978-3-942036-33-7
- 230 Faigle, Benjamin: *Adaptive modelling of compositional multi-phase flow with capillary pressure*, 2014, ISBN 978-3-942036-34-4
- 231 Oladyshkin, Sergey: *Efficient modeling of environmental systems in the face of complexity and uncertainty*, 2014, ISBN 978-3-942036-35-1
- 232 Sugimoto, Takayuki: *Copula based Stochastic Analysis of Discharge Time Series*, 2014, ISBN 978-3-942036-36-8
- 233 Koch, Jonas: *Simulation, Identification and Characterization of Contaminant Source Architectures in the Subsurface*, 2014, ISBN 978-3-942036-37-5
- 234 Zhang, Jin: *Investigations on Urban River Regulation and Ecological Rehabilitation Measures, Case of Shenzhen in China*, 2014, ISBN 978-3-942036-38-2
- 235 Siebel, Rüdiger: *Experimentelle Untersuchungen zur hydrodynamischen Belastung und Standsicherheit von Deckwerken an überströmbaren Erddämmen*, 2014, ISBN 978-3-942036-39-9

- 236 Baber, Katherina: *Coupling free flow and flow in porous media in biological and technical applications: From a simple to a complex interface description*, 2014, ISBN 978-3-942036-40-5
- 237 Nuske, Klaus Philipp: *Beyond Local Equilibrium — Relaxing local equilibrium assumptions in multiphase flow in porous media*, 2014, ISBN 978-3-942036-41-2
- 238 Geiges, Andreas: *Efficient concepts for optimal experimental design in nonlinear environmental systems*, 2014, ISBN 978-3-942036-42-9
- 239 Schwenck, Nicolas: *An XFEM-Based Model for Fluid Flow in Fractured Porous Media*, 2014, ISBN 978-3-942036-43-6
- 240 Chamorro Chávez, Alejandro: *Stochastic and hydrological modelling for climate change prediction in the Lima region, Peru*, 2015, ISBN 978-3-942036-44-3
- 241 Yulizar: *Investigation of Changes in Hydro-Meteorological Time Series Using a Depth-Based Approach*, 2015, ISBN 978-3-942036-45-0
- 242 Kretschmer, Nicole: *Impacts of the existing water allocation scheme on the Limarí watershed – Chile, an integrative approach*, 2015, ISBN 978-3-942036-46-7
- 243 Kramer, Matthias: *Luftbedarf von Freistrahlturbinen im Gegendruckbetrieb*, 2015, ISBN 978-3-942036-47-4
- 244 Hommel, Johannes: *Modeling biogeochemical and mass transport processes in the subsurface: Investigation of microbially induced calcite precipitation*, 2016, ISBN 978-3-942036-48-1
- 245 Germer, Kai: *Wasserinfiltration in die ungesättigte Zone eines makroporösen Hanges und deren Einfluss auf die Hangstabilität*, 2016, ISBN 978-3-942036-49-8
- 246 Hörning, Sebastian: *Process-oriented modeling of spatial random fields using copulas*, 2016, ISBN 978-3-942036-50-4
- 247 Jambhekar, Vishal: *Numerical modeling and analysis of evaporative salinization in a coupled free-flow porous-media system*, 2016, ISBN 978-3-942036-51-1
- 248 Huang, Yingchun: *Study on the spatial and temporal transferability of conceptual hydrological models*, 2016, ISBN 978-3-942036-52-8
- 249 Kleinknecht, Simon Matthias: *Migration and retention of a heavy NAPL vapor and remediation of the unsaturated zone*, 2016, ISBN 978-3-942036-53-5
- 250 Kwakye, Stephen Oppong: *Study on the effects of climate change on the hydrology of the West African sub-region*, 2016, ISBN 978-3-942036-54-2
- 251 Kissinger, Alexander: *Basin-Scale Site Screening and Investigation of Possible Impacts of CO₂ Storage on Subsurface Hydrosystems*, 2016, ISBN 978-3-942036-55-9
- 252 Thomas Müller, *Generation of a Realistic Temporal Structure of Synthetic Precipitation Time Series for Sewer Applications*, 2017, ISBN 978-3-942036-56-6

Die Mitteilungshefte ab der Nr. 134 (Jg. 2005) stehen als pdf-Datei über die Homepage des Instituts: www.iws.uni-stuttgart.de zur Verfügung.